

**SEARCH REQUEST FORM**

Scientific and Technical Information Center

Requester's Full Name \_\_\_\_\_ Examiner # : \_\_\_\_\_ Date: \_\_\_\_\_  
 Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
 Mail Box and Bldg Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*

Please provide a statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or priority of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

**STAFF USE ONLY**

Type of Search		Vendors and cost where applicable
Searcher <u>BA Beverly C4994</u>	NA Sequence (#) _____	STN _____
Searcher Phone # _____	AA Sequence (#) _____	Dialog _____
Searcher Location _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Filled _____	Bibliographic _____	Dr Link _____
Date Completed <u>02-25-00</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time _____	Fulltext _____	Sequence Systems _____
Clerical Prep Time _____	Patent Family _____	WWW Internet _____
Online Time _____	Other _____	Other (specify) <u>MP</u>

The first part of the paper discusses the importance of the study of the history of the United States. It is argued that the study of history is essential for a full understanding of the present. The second part of the paper discusses the importance of the study of the history of the United States. It is argued that the study of history is essential for a full understanding of the present. The third part of the paper discusses the importance of the study of the history of the United States. It is argued that the study of history is essential for a full understanding of the present.

Bovin  
09/132799  
508 1251  
u/l

\*\*\*\*\*

(三)

一、二、三、四、五、六、七、八、九、十

**Bull. Environ. Contam. Toxicol.**

Department of Chemistry, University of Edinburgh, U.K.

© 1997 by the American Society of Cell Biology. All rights reserved. No part of this publication may be reproduced without permission from the American Society of Cell Biology.

measured using protein-protein database search, using Smith Waterman algorithm

[illegible]

44.736 Million cell updates/sec

Downloaded from www.worldscientific.com by UNIVERSITY OF NEWCASTLE on 07/06/18. For personal use only.

Political Science 4  
 Government 4  
 VPR 4

5011-10-01; 152453997 residues

### Post-Processing: Minimum Match 90%

List for first 1000 summaries

[illegible]

1985	12:07	13:08	14:08	15:06	16:09	17:09
1986	12:06	13:07	14:07	15:05	16:08	17:08
1987	12:05	13:06	14:06	15:04	16:07	17:07
1988	12:04	13:05	14:05	15:03	16:06	17:06
1989	12:03	13:04	14:04	15:02	16:05	17:05
1990	12:02	13:03	14:03	15:01	16:04	17:04
1991	12:01	13:02	14:02	15:00	16:03	17:03
1992	12:00	13:01	14:01	15:00	16:02	17:02

$$\begin{aligned} & \mathcal{M}(\mathcal{C}) \cap \mathcal{M}(\mathcal{C}') = \mathcal{M}(\mathcal{C} \cap \mathcal{C}') \\ & \mathcal{M}(\mathcal{C}) \cup \mathcal{M}(\mathcal{C}') = \mathcal{M}(\mathcal{C} \cup \mathcal{C}') \\ & \mathcal{M}(\mathcal{C}) \cap \mathcal{M}(\mathcal{C}') \cap \mathcal{M}(\mathcal{C}'') = \mathcal{M}(\mathcal{C} \cap \mathcal{C}' \cap \mathcal{C}'') \\ & \mathcal{M}(\mathcal{C}) \cup \mathcal{M}(\mathcal{C}') \cup \mathcal{M}(\mathcal{C}'') = \mathcal{M}(\mathcal{C} \cup \mathcal{C}' \cup \mathcal{C}'') \end{aligned}$$

Statistics. Mean 14.48, Variance 20.47, scale 0.59

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed

and is derived by analysis of the total score distribution.

## SUMMARY

Ref.	Year	Match	Length	HR	FD	Description	Pred. No.
1	1960	1	10	10	10	10	10
2	1961	2	20	20	20	20	20
3	1962	3	30	30	30	30	30
4	1963	4	40	40	40	40	40
5	1964	5	50	50	50	50	50
6	1965	6	60	60	60	60	60
7	1966	7	70	70	70	70	70
8	1967	8	80	80	80	80	80
9	1968	9	90	90	90	90	90
10	1969	10	100	100	100	100	100
11	1970	11	110	110	110	110	110
12	1971	12	120	120	120	120	120
13	1972	13	130	130	130	130	130
14	1973	14	140	140	140	140	140
15	1974	15	150	150	150	150	150
16	1975	16	160	160	160	160	160
17	1976	17	170	170	170	170	170
18	1977	18	180	180	180	180	180
19	1978	19	190	190	190	190	190
20	1979	20	200	200	200	200	200
21	1980	21	210	210	210	210	210
22	1981	22	220	220	220	220	220
23	1982	23	230	230	230	230	230
24	1983	24	240	240	240	240	240
25	1984	25	250	250	250	250	250
26	1985	26	260	260	260	260	260
27	1986	27	270	270	270	270	270
28	1987	28	280	280	280	280	280
29	1988	29	290	290	290	290	290
30	1989	30	300	300	300	300	300
31	1990	31	310	310	310	310	310
32	1991	32	320	320	320	320	320
33	1992	33	330	330	330	330	330
34	1993	34	340	340	340	340	340
35	1994	35	350	350	350	350	350
36	1995	36	360	360	360	360	360
37	1996	37	370	370	370	370	370
38	1997	38	380	380	380	380	380
39	1998	39	390	390	390	390	390
40	1999	40	400	400	400	400	400
41	2000	41	410	410	410	410	410
42	2001	42	420	420	420	420	420
43	2002	43	430	430	430	430	430
44	2003	44	440	440	440	440	440
45	2004	45	450	450	450	450	450
46	2005	46	460	460	460	460	460
47	2006	47	470	470	470	470	470
48	2007	48	480	480	480	480	480
49	2008	49	490	490	490	490	490
50	2009	50	500	500	500	500	500
51	2010	51	510	510	510	510	510
52	2011	52	520	520	520	520	520
53	2012						

Case	US-09-142-	Sequence 1, Applicatio	2,980+03
41	101,0		
42	101,0		
43	101,0		
44	101,0		
45	101,0		
46	101,0		
47	101,0		
48	101,0		
49	101,0		
50	101,0		
51	101,0		
52	101,0		
53	101,0		
54	101,0		
55	101,0		
56	101,0		
57	101,0		
58	101,0		
59	101,0		
60	101,0		
61	101,0		
62	101,0		
63	101,0		
64	101,0		
65	101,0		
66	101,0		
67	101,0		
68	101,0		
69	101,0		
70	101,0		
71	101,0		
72	101,0		
73	101,0		
74	101,0		
75	101,0		
76	101,0		
77	101,0		
78	101,0		
79	101,0		
80	101,0		
81	101,0		
82	101,0		
83	101,0		
84	101,0		
85	101,0		
86	101,0		
87	101,0		
88	101,0		
89	101,0		
90	101,0		
91	101,0		
92	101,0		
93	101,0		
94	101,0		
95	101,0		
96	101,0		
97	101,0		
98	101,0		
99	101,0		
100	101,0		

Sequence	Applicatio	2.986+03
Sequence 2	Applicatio	2.986+03
Sequence 3	Applicatio	2.986+03

[illegible]

Run	Time	Sequence	Applicator	Conc.
1	15	MS-09-010	22	2.980-03
2	15	MS-09-010	22	2.980-03
3	15	MS-09-010	22	2.980-03
4	15	MS-09-010	22	2.980-03
5	15	MS-09-010	22	2.980-03
6	15	MS-09-010	22	2.980-03
7	15	MS-09-010	22	2.980-03
8	15	MS-09-010	22	2.980-03
9	15	MS-09-010	22	2.980-03
10	15	MS-09-010	22	2.980-03
11	15	MS-09-010	22	2.980-03
12	15	MS-09-010	22	2.980-03
13	15	MS-09-010	22	2.980-03
14	15	MS-09-010	22	2.980-03
15	15	MS-09-010	22	2.980-03
16	15	MS-09-010	22	2.980-03
17	15	MS-09-010	22	2.980-03
18	15	MS-09-010	22	2.980-03
19	15	MS-09-010	22	2.980-03
20	15	MS-09-010	22	2.980-03
21	15	MS-09-010	22	2.980-03
22	15	MS-09-010	22	2.980-03
23	15	MS-09-010	22	2.980-03
24	15	MS-09-010	22	2.980-03
25	15	MS-09-010	22	2.980-03
26	15	MS-09-010	22	2.980-03
27	15	MS-09-010	22	2.980-03
28	15	MS-09-010	22	2.980-03
29	15	MS-09-010	22	2.980-03
30	15	MS-09-010	22	2.980-03
31	15	MS-09-010	22	2.980-03
32	15	MS-09-010	22	2.980-03
33	15	MS-09-010	22	2.980-03
34	15	MS-09-010	22	2.980-03
35	15	MS-09-010	22	2.980-03
36	15	MS-09-010	22	2.980-03
37	15	MS-09-010	22	2.980-03
38	15	MS-09-010	22	2.980-03
39	15	MS-09-010	22	2.980-03
40	15	MS-09-010	22	2.980-03
41	15	MS-09-010	22	2.980-03
42	15	MS-09-010	22	2.980-03
43	15	MS-09-010	22	2.980-03
44	15	MS-09-010	22	2.980-03
45	15	MS-09-010	22	2.980-03
46	15	MS-09-010	22	2.980-03
47	15	MS-09-010	22	2.980-03
48	15	MS-09-010	22	2.980-03
49	15	MS-09-010	22	2.980-03
50	15	MS-09-010	22	2.980-03
51	15	MS-09-010	22	2.980-03
52	15	MS-09-010	22	2.980-03
53	15	MS-09-010	22	2.980-03
54	15	MS-09-010	22	2.980-03
55	15	MS-09-010	22	2.980-03
56	15	MS-09-010	22	2.980-03
57	15	MS-09-010	22	2.980-03
58	15	MS-09-010	22	2.980-03
59	15	MS-09-010	22	2.980-03
60	15	MS-09-010	22	2.980-03
61	15	MS-09-010	22	2.980-03
62	15	MS-09-010	22	2.980-03
63	15	MS-09-010	22	2.980-03
64	15	MS-09-010	22	2.980-03
65	15	MS-09-010	22	2.980-03
66	15	MS-09-010	22	2.980-03
67	15	MS-09-010	22	2.980-03
68	15	MS-09-010	22	2.980-03
69	15	MS-09-010		

Sequence	Sequence 3, Applied	Sequence 3, Applied	Sequence 3, Applied
1	100.0	100.0	100.0
2	100.0	100.0	100.0
3	100.0	100.0	100.0
4	100.0	100.0	100.0
5	100.0	100.0	100.0
6	100.0	100.0	100.0
7	100.0	100.0	100.0
8	100.0	100.0	100.0
9	100.0	100.0	100.0
10	100.0	100.0	100.0
11	100.0	100.0	100.0
12	100.0	100.0	100.0
13	100.0	100.0	100.0
14	100.0	100.0	100.0
15	100.0	100.0	100.0
16	100.0	100.0	100.0
17	100.0	100.0	100.0
18	100.0	100.0	100.0
19	100.0	100.0	100.0
20	100.0	100.0	100.0
21	100.0	100.0	100.0
22	100.0	100.0	100.0
23	100.0	100.0	100.0
24	100.0	100.0	100.0
25	100.0	100.0	100.0
26	100.0	100.0	100.0
27	100.0	100.0	100.0
28	100.0	100.0	100.0
29	100.0	100.0	100.0
30	100.0	100.0	100.0
31	100.0	100.0	100.0
32	100.0	100.0	100.0
33	100.0	100.0	100.0
34	100.0	100.0	100.0
35	100.0	100.0	100.0
36	100.0	100.0	100.0
37	100.0	100.0	100.0
38	100.0	100.0	100.0
39	100.0	100.0	100.0
40	100.0	100.0	100.0
41	100.0	100.0	100.0
42	100.0	100.0	100.0
43	100.0	100.0	100.0
44	100.0	100.0	100.0
45	100.0	100.0	100.0
46	100.0	100.0	100.0
47	100.0	100.0	100.0
48	100.0	100.0	100.0
49	100.0	100.0	100.0
50	100.0	100.0	100.0
51	100.0	100.0	100.0
52	100.0	100.0	100.0
53	100.0	100.0	100.0
54	100.0	100.0	100.0
55	100.0	100.0	100.0
56	100.0	100.0	100.0
57	100.0	100.0	100.0
58	100.0	100.0	100.0
59	100.0	100.0	100.0
60	100.0	100.0	100.0
61	100.0	100.0	100.0
62	100.0	100.0	100.0
63	100.0	100.0	100.0
64	100.0	100.0	100.0
65	100.0	100.0	100.0
66	100.0	100.0	100.0
67	100.0	100.0	100.0
68	100.0	100.0	100.0
69	100.0	100.0	100.0
70	100.0	100.0	100.0
71	100.0	100.0	100.0
72	100.0	100.0	100.0
73	100.0	100.0	100.0
74	100.0	100.0	100.0
75	100.0	100.0	100.0
76	100.0	100.0	100.0
77	100.0	100.0	100.0
78	100.0	100.0	100.0
79	100.0	100.0	100.0
80	100.0	100.0	100.0
81	100.0	100.0	100.0
82	100.0	100.0	100.0
83	100.0	100.0	100.0
84	100.0	100.0	100.0
85			

Year	Volume	Issue	Page	Author	Year	Volume	Issue	Page	Author
1990	1	1	1	1	1990	1	1	1	1
1991	2	2	2	2	1991	2	2	2	2
1992	3	3	3	3	1992	3	3	3	3
1993	4	4	4	4	1993	4	4	4	4
1994	5	5	5	5	1994	5	5	5	5
1995	6	6	6	6	1995	6	6	6	6
1996	7	7	7	7	1996	7	7	7	7
1997	8	8	8	8	1997	8	8	8	8
1998	9	9	9	9	1998	9	9	9	9
1999	10	10	10	10	1999	10	10	10	10
2000	11	11	11	11	2000	11	11	11	11
2001	12	12	12	12	2001	12	12	12	12
2002	13	13	13	13	2002	13	13	13	13
2003	14	14	14	14	2003	14	14	14	14
2004	15	15	15	15	2004	15	15	15	15
2005	16	16	16	16	2005	16	16	16	16
2006	17	17	17	17	2006	17	17	17	17
2007	18	18	18	18	2007	18	18	18	18
2008	19	19	19	19	2008	19	19	19	19
2009	20	20	20	20	2009	20	20	20	20
2010	21	21	21	21	2010	21	21	21	21
2011	22	22	22	22	2011	22	22	22	22
2012	23	23	23	23	2012	23	23	23	23
2013	24	24	24	24	2013	24	24	24	24
2014	25	25	25	25	2014	25	25	25	25
2015	26	26	26	26	2015	26	26	26	26
2016	27	27	27	27	2016	27	27	27	27
2017	28	28	28	28	2017	28	28	28	28
2018	29	29	29	29	2018	29	29	29	29
2019	30	30	30	30	2019	30	30	30	30
2020	31	31	31	31	2020	31	31	31	31
2021	32	32	32	32	2021	32	32	32	32
2022	33	33	33	33	2022	33	33	33	33
2023	34	34	34	34	2023	34	34	34	34
2024	35	35	35	35	2024	35	35	35	35
2025	36	36	36	36	2025	36	36	36	36
2026	37	37	37	37	2026	37	37	37	37
2027	38	38	38	38	2027	38	38	38	38
2028	39	39	39	39	2028	39	39	39	39
2029	40	40	40	40	2029	40	40	40	40
2030	41	41	41	41	2030	41	41	41	41
2031	42	42	42	42	2031	42	42	42	42
2032	43	43	43	43	2032	43	43	43	43
2033	44	44	44	44	2033	44	44	44	44
2034	45	45	45	45	2034	45	45		

94	POT-US99-02	sequence 97, Applicant	2,996.05
95	POT-US99-03	sequence 97, Applicant	2,989.05

43	2	75-60-167-	Sequence 1702, Applica	2,988-03
43	2	75-60-167-	Sequence 1466, Applic	2,988-03

[illegible][illegible]

US-60-100	sequence 4927	Applica	2,980+03
US-60-109	sequence 4930	Applica	2,980+03

97	US-60-164-	Sequence 1070, Applica	2,980.03
97	US-60-173-	Sequence 1922, Affiliat	2,990.03

Journal of Management Inquiry 25(4) 391-407 © The Author(s) 2016. Reprints and permissions: [sagepub.com/journalsPermissions.nav](http://sagepub.com/journalsPermissions.nav) DOI: 10.1177/1056492616666666









[illegible]

[illegible]



```

971 24 92 5 720 2 US-60-150- Sequence 482, Applicat 5.84e+03
972 24 92 5 721 2 US-60-152- Sequence 1521, Applic 5.84e+03
973 24 92 5 722 2 US-60-153- Sequence 4023, Applic 5.84e+03
974 24 92 5 723 2 US-60-154- Sequence 5, Applic 5.84e+03
975 24 92 5 724 2 US-60-155- Sequence 8421, Applic 5.84e+03
976 24 92 5 725 2 US-60-156- Sequence 2622, Applic 5.84e+03
977 24 92 5 726 2 US-60-157- Sequence 15313, Applic 5.84e+03
978 24 92 5 727 2 US-60-158- Sequence 20513, Applic 5.84e+03
979 24 92 5 728 2 US-60-159- Sequence 0018, Applic 5.84e+03
980 24 92 5 729 2 US-60-160- Sequence 651, Applic 5.84e+03
981 24 92 5 730 2 US-60-161- Sequence 11231, Applic 5.84e+03
982 24 92 5 731 2 US-60-162- Sequence 1522, Applic 5.84e+03
983 24 92 5 732 2 US-60-163- Sequence 59116, Applic 5.84e+03
984 24 92 5 733 2 US-60-164- Sequence 6197, Applic 5.84e+03
985 24 92 5 734 2 US-60-165- Sequence 2010, Applic 5.84e+03
986 24 92 5 735 2 US-60-166- Sequence 202, Applic 5.84e+03
987 24 92 5 736 2 US-60-167- Sequence 4254, Applic 5.84e+03
988 24 92 5 737 2 US-60-168- Sequence 18229, Applic 5.84e+03
989 24 92 5 738 2 US-60-169- Sequence 3803, Applic 5.84e+03
990 24 92 5 739 2 US-60-170- Sequence 1253, Applic 5.84e+03
991 24 92 5 740 2 US-60-171- Sequence 1457, Applic 5.84e+03
992 24 92 5 741 2 US-60-172- Sequence 1479, Applic 5.84e+03
993 24 92 5 742 2 US-60-173- Sequence 10045, Applic 5.84e+03
994 24 92 5 743 2 US-60-174- Sequence 184, Applic 5.84e+03
995 24 92 5 744 2 US-60-175- Sequence 176, Applic 5.84e+03
996 24 92 5 745 2 US-60-176- Sequence 5189, Applic 5.84e+03
997 24 92 5 746 2 US-60-177- Sequence 18533, Applic 5.84e+03
998 24 92 5 747 2 US-60-178- Sequence 2250, Applic 5.84e+03
999 24 92 5 748 2 US-60-179- Sequence 1752, Applic 5.84e+03
1000 24 92 5 749 2 US-60-180- Sequence 1752, Applic 5.84e+03

```

## ALIGNMENTS

```

RESULT 1
ID US-09-132-799-1 STANFAPD: PPT: 4 AA.
XX AC xxxxxx
DE Sequence 1, Application US/09132799
DE Sequence 1, Application US/09132799
DE GENERAL INFORMATION:
DE APPLICANT: Schonrock, Uwe
DE APPLICANT: Max, Heiner
DE TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL PREPARATIONS COMPRISING
DE TITLE OF INVENTION: OLIGOPEPTIDES FOR LIGHTENING THE SKIN OF AGE MARKS
DE TITLE OF INVENTION: AND/OR FOR PREVENTING TANNING OF THE SKIN, IN
DE FILE REFERENCE: Patsdorff-514-KGB
DE CURRENT APPLICATION NUMBER: US/09/132-799
DE CURRENT FILING DATE: 1998-08-13
DE NUMBER OF SEQ ID NOS: 3
DE SOFTWARE: PatentIn Ver. 2.1
DE SEQ ID NO 1
DE LENGTH: 4
DE TYPE: PPT
DE ORGANISM: Artificial Sequence
DE FEATURE:
DE OTHER INFORMATION: Description of Artificial Sequence: structural
DE OTHER INFORMATION: base for oligopeptides
SQ SEQUENCE 4 AA: 470 MW: 126.0N

Query Match 100.0% Score 31, DB 16, Length 4,
Best Local Similarity 100.0%, Prod. No. 2.98e-03,
Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 1 VWRP 4
III
CY 1 VWRP 4

```

```

RESULT 2
ID US-09-132-799-2 STANFAPD: PPT: 5 AA.
XX AC xxxxxx
DE Sequence 2, Application US/09132799
DE Sequence 2, Application US/09132799
DE GENERAL INFORMATION:
DE APPLICANT: Schonrock, Uwe
DE APPLICANT: Max, Heiner
DE TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL PREPARATIONS COMPRISING
DE TITLE OF INVENTION: OLIGOPEPTIDES FOR LIGHTENING THE SKIN OF AGE MARKS
DE TITLE OF INVENTION: AND/OR FOR PREVENTING TANNING OF THE SKIN, IN
DE FILE REFERENCE: Patsdorff-514-KGB
DE CURRENT APPLICATION NUMBER: US/09/132-799
DE CURRENT FILING DATE: 1998-08-13
DE NUMBER OF SEQ ID NOS: 3
DE SOFTWARE: PatentIn Ver. 2.1
DE SEQ ID NO 2
DE LENGTH: 5
DE TYPE: PPT
DE ORGANISM: Artificial Sequence
DE FEATURE:
DE OTHER INFORMATION: Description of Artificial Sequence: structural base
DE OTHER INFORMATION: for oligopeptides
SQ SEQUENCE 5 AA: 567 MW: 201.0N

Query Match 100.0% Score 31, DB 16, Length 5;
Best Local Similarity 100.0%, Prod. No. 2.98e-03,
Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 1 VWRP 4
III
CY 1 VWRP 4

```

```

RESULT 3
ID US-09-132-799-3 STANFAPD: PPT: 6 AA.
XX AC xxxxxx
DE Sequence 3, Application US/09132799
DE Sequence 3, Application US/09132799
DE GENERAL INFORMATION:
DE APPLICANT: Schonrock, Uwe
DE APPLICANT: Max, Heiner
DE TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL PREPARATIONS COMPRISING
DE TITLE OF INVENTION: OLIGOPEPTIDES FOR LIGHTENING THE SKIN OF AGE MARKS
DE TITLE OF INVENTION: AND/OR FOR PREVENTING TANNING OF THE SKIN, IN
DE FILE REFERENCE: Patsdorff-514-KGB
DE CURRENT APPLICATION NUMBER: US/09/132-799
DE CURRENT FILING DATE: 1998-08-13
DE NUMBER OF SEQ ID NOS: 3
DE SOFTWARE: PatentIn Ver. 2.1
DE SEQ ID NO 3
DE LENGTH: 6
DE TYPE: PPT
DE ORGANISM: Artificial Sequence
DE FEATURE:
DE OTHER INFORMATION: Description of Artificial Sequence: Structural
DE OTHER INFORMATION: base for oligopeptides.

```



CC GENERAL INFORMATION:  
 CC APPLICANT: Schatz, Peter J.  
 CC APPLICANT: Cull, Millard G.  
 CC APPLICANT: Miller, Jeff F.  
 CC APPLICANT: Stemmer, William P.C.  
 CC APPLICANT: Gates, Christian M.  
 CC TITLE OF INVENTION: Peptide Library and Screening Method  
 CC NUMBER OF SEQUENCES: 152  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: William M. Smith  
 CC STREET: One Market Plaza, Stewart Tower, Suite 2000  
 CC CITY: San Francisco  
 CC STATE: California  
 CC COUNTRY: USA  
 CC ZIP: 94104  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.20  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/09/010,216  
 CC FILING DATE:  
 CC CLASSIFICATION:  
 CC PRACTICE APPLICATION DATA:  
 CC APPLICATION NUMBER: US 08,549,341  
 CC FILING DATE: 26-OCT-1995  
 CC APPLICATION NUMBER: US 08,129,641  
 CC FILING DATE: 15 AUG-1994  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 07/963,121  
 CC FILING DATE: 15-OCT-1992  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Smith, William M.  
 CC REGISTRATION NUMBER: 30,723  
 CC REFERENCE/DOCKET NUMBER: 15,000,0124-05  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 415-926-2400  
 CC TELEFAX: 415-926-2422  
 CC INFORMATION FOR SEQ ID NO. 32  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 12 amino acids  
 CC TYPE: amino acid  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: peptide  
 CC IMMEDIATE SOURCE:  
 CC CLONE: 59-2-1-1  
 CC SEQUENCE 12 AA: 1376 MW: 912 CH:  
 CC Query Match 100.0% Score 31: DB 15: Length 12:  
 CC Best Local Similarity 100.0% Prid No. 4 348-03  
 CC Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 8 VWRP 11  
 Q3 1 VWRP 4

RESULT 7  
 ID US-08-622-632-33 STANDARD: PRI: 15 AA:  
 XX AC xxxxxx  
 XX DI  
 XX SEQUENCE 33: Application US/08/010,216  
 XX GENERAL INFORMATION:  
 XX APPLICANT: Anthony C. Forster  
 XX TITLE OF INVENTION: Immunologic Screening Process, and  
 XX TITLE OF INVENTION: Compositions Thereof

CC NUMBER OF SEQUENCES: 106  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: LAHIVE & COOKFIELD  
 CC STREET: 60 State Street  
 CC CITY: Boston  
 CC STATE: Massachusetts  
 CC COUNTRY: USA  
 CC ZIP: 02109-1975  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/010,216  
 CC FILING DATE: 21-MARCH-1996  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER:  
 CC FILING DATE:  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Vincent, Matthew P.  
 CC REGISTRATION NUMBER: 30,716  
 CC REFERENCE/DOCKET NUMBER: HMT-018  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 1617-227-7400  
 CC TELEFAX: 1617-227-5241  
 CC INFORMATION FOR SEQ ID NO. 33:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 15 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: peptide  
 CC FRAGMENT TYPE: internal  
 CC SEQUENCE 15 AA: 1502 MW: 1580 CH:  
 CC Query Match 100.0% Score 31: DB 15: Length 15:  
 CC Best Local Similarity 100.0% Prid No. 2 980-03:  
 CC Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Pe 8 VWRP 11  
 Qy 1 VWRP 4

RESULT 8  
 ID US-08-488-315A-8 STANDARD: PRI: 15 AA:  
 XX AC xxxxxx  
 XX DI  
 XX SEQUENCE 8: Application US/08/010,216  
 XX GENERAL INFORMATION:  
 XX APPLICANT: Atassi, M. Zouhair  
 XX TITLE OF INVENTION: Detection and Treatment of Myasthenia  
 XX TITLE OF INVENTION: Graves  
 XX NUMBER OF SEQUENCES: 25  
 XX CORRESPONDENCE ADDRESS:  
 XX ADDRESSEE: Stephen McDaniel  
 XX STREET: P.O. Box 3267  
 XX CITY: Houston  
 XX STATE: Texas  
 XX COUNTRY: USA  
 XX ZIP: 77052-3267  
 XX COMPUTER READABLE FORM:  
 XX MEDIUM TYPE: Floppy disk  
 XX COMPUTER: IBM PC compatible  
 XX OPERATING SYSTEM: MS-DOS  
 XX SOFTWARE: WordPerfect 5.1  
 XX CURRENT APPLICATION DATA:  
 XX APPLICATION NUMBER: US/08/488,315A























CC TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,  
 CC TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES  
 CC TITLE OF INVENTION: THEREOF  
 CC FILE REFERENCE: CL000173  
 CC CURRENT APPLICATION NUMBER: US/60/173,404  
 CC CURRENT FILING DATE: 1999-12-29  
 CC NUMBER OF SEQ ID NOS: 30269  
 CC SOFTWARE: FastSeq for Windows Version 4.0  
 CC SEQ ID NO 12740  
 CC LENGTH: 196  
 CC TYPE: PRI  
 CC ORGANISM: Drosophila  
 SQ SEQUENCE 196 AA: 22128 MW: 190309 CN:

Query Match: 100.0% Score 39 DB 2 Length 196;  
 Best Local Similarity 100.0% Pred. No. 6,14e-02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 188 VWRPP 192  
 |||||  
 QV 1 VWRPP 5

RESULT 15

ID US-60-167-217-15541 STANDARD; PRI: 196 AA.

AC XXXXX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Sequence 15541, Application US/60167217

Sequence 15541, Application US/60167217

GENERAL INFORMATION:

APPLICANT: LI, Peter W. D.

TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC

TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES

TITLE OF INVENTION: THEREOF

FILE REFERENCE: CL000152

CURRENT APPLICATION NUMBER: US/60/173,404

CURRENT FILING DATE: 1999-11-24

NUMBER OF SEQ ID NOS: 33195

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 15541

LENGTH: 196

TYPE: PRI

ORGANISM: Drosophila

SEQUENCE 195 AA: 22128 MW: 190309 CN:

Query Match: 100.0% Score 39 DB 2 Length 196;  
 Best Local Similarity 100.0% Pred. No. 6,14e-02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 188 VWRPP 192  
 |||||  
 QV 1 VWRPP 5

Search completed: Fri Feb 25 13:23:49 2000  
 Job time : 229 secs.



CC FILING DATE: 23 JUN 1995  
 CC CLASSIFICATION: S14  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Spratt, Gwendolyn D.  
 CC REGISTRATION NUMBER: 36,016  
 CC REFERENCE/KEY NUMBER: 14014 0135  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (404) 688-0770  
 CC TELEFAX: (404) 688-9880  
 CC INFORMATION FOR SEQ ID NO: 2:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 131 amino acids  
 CC TYPE: amino acid  
 CC STRANDEDNESS: not relevant  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: peptide  
 CC STRANDEDNESS: not relevant  
 CC TOPOLOGY: linear  
 CC SEQUENCE: 131 AA: 14255 MW: 84197 CN:  
 Query Match: 100.0% Score 47 DB 14 Length 131  
 Best Local Similarity 100.0% Prod No 1 96e-02  
 Matches 6 Conservative 0 Mismatches 0 Indels 0 Gaps 0  
 DB 83 VRPPP 88  
 QY 1 VRPPP 6  
 RESULT 2  
 ID US-09-000-436-2 STANDARD: PPT: 131 AA  
 XX xxxxxx  
 CC SEQUENCE 2: Application US/50000436  
 CC SCHEME 2: Application US/50000436  
 CC GENERAL INFORMATION:  
 CC APPLICANT: HEARING, Vincent J., Jr.  
 CC TITLE OF INVENTION: Depigmentation Activity of Aquatic Signal  
 CC TITLE OF INVENTION: Protein and Peptides Thereof  
 CC NUMBER OF SEQUENCES: 29  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Needle & Posenberg, P.C.  
 CC STREET: Suite 1200, 127 Peachtree St., N.E.  
 CC CITY: Atlanta  
 CC STATE: GA  
 CC COUNTRY: USA  
 CC ZIP: 30303-1811  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patent In Release #1.0, Version #1.30  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/90/000,436  
 CC FILING DATE: 22 Dec 1997  
 CC CLASSIFICATION:  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: PCT/US96/10695  
 CC FILING DATE: 21 Jun 1996  
 CC CLASSIFICATION:  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: 602,055,436  
 CC FILING DATE: 23 Jun 1995  
 CC CLASSIFICATION:  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Spratt, Gwendolyn D.  
 CC REGISTRATION NUMBER: 36,016  
 CC REFERENCE/KEY NUMBER: 14014 0135  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (404) 688-0770  
 CC TELEFAX: (404) 688-9880

CC INFORMATION FOR SEQ ID NO: 2:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 131 amino acids  
 CC TYPE: amino acid  
 CC STRANDEDNESS: not relevant  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: peptide  
 CC SEQUENCE: 131 AA: 14255 MW: 84197 CN:  
 Query Match: 100.0% Score 47 DB 20 Length 131  
 Best Local Similarity 100.0% Prod No 1 96e-02  
 Matches 6 Conservative 0 Mismatches 0 Indels 0 Gaps 0  
 DB 83 VRPPP 88  
 QY 1 VRPPP 6  
 RESULT 4  
 ID US-09-417-507-200071 STANDARD: PPT: 70 AA  
 XX xxxxxx  
 CC SEQUENCE 20071: Application US/00417507  
 CC SCHEME 22971: Application US/00417507  
 CC GENERAL INFORMATION:  
 CC APPLICANT: KEITH G. WEINSTOCK ET AL.  
 CC TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILL  
 CC TITLE OF INVENTION: FUNGICIDES FOR DIAGNOSTICS AND THERAPEUTICS  
 CC FILE REFERENCE: PATH99-10  
 CC CURRENT APPLICATION NUMBER: US/00/417,507  
 CC CURRENT FILING DATE: 1999-10-14  
 CC SEQ ID NO 20471  
 CC LENGTH: 70  
 CC TYPE: PPT  
 CC ORGANISM: A.fumigatus  
 CC SEQUENCE: 70 AA: 8081 MW: 24553 CN:  
 Query Match: 95.7% Score 45 DB 19 Length 70  
 Best Local Similarity 93.9% Prod No 1 32e-02  
 Matches 5 Conservative 1 Mismatches 0 Indels 0 Gaps 0  
 DB 19 VRPPP 24  
 QY 1 VRPPP 6  
 RESULT 5  
 ID US-60-174-464-30179 STANDARD: PPT: 216 AA  
 XX xxxxxx  
 CC SEQUENCE 30179: Application US/50173464  
 CC SCHEME 30179: Application US/50173464  
 CC GENERAL INFORMATION:  
 CC APPLICANT: LI, Peter W.D.  
 CC TITLE OF INVENTION: ISOLATED PROTEIN ENDED RECEPTORS,  
 CC TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING SUCH PROTEINS AND USES  
 CC TITLE OF INVENTION: THEREOF  
 CC FILE REFERENCE: 01090173  
 CC CURRENT APPLICATION NUMBER: US/00/174,464  
 CC CURRENT FILING DATE: 1999-12-29  
 CC NUMBER OF SEQ ID NOS: 30269  
 CC SOFTWARE: FASTSEQ for Windows Version 4.0  
 CC SEQ ID NO 30179  
 CC LENGTH: 216

```

11 TYPE: PRT
12 ORGANISM: Drosophila
13 SEQUENCE 415 AA: 2354 MW: 24137.0N.

Query Match
  Best Local Similarity: 95.7% Score 45: DB 2: Length 216:
  Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

14 199 VRRPP 214
15 1 VVRRPP 6

RESULT 8
16 US-60-161-942-164 STANDARD: PRT: 415 AA.
17 AC XXXXX
18 DT
19 DE
20 XX
21 XX
22 XX
23 SEQUENCE 1964, Application US/60161942
24 SEQUENCE 1964, Application US/60161942
25 GENERAL INFORMATION:
26 APPLICANT: Venter, J. Craig
27 TITLE OF INVENTION: The Drosophila Genome, Primary Nucleic
28 TITLE OF INVENTION: Acid Sequences, Systems Containing the Nucleic Acid Sequen
29 TITLE OF INVENTION: and Uses Thereof
30 FILE REFERENCE: 1600122
31 CURRENT APPLICATION NUMBER: US/60161942
32 CURRENT FILING DATE: 1999-10-28
33 NUMBER OF SEQ ID NOS: 2620
34 SOFTWARE: FASTSEQ for Windows Version 4.0
35 SEQ ID NO: 1964
36 LENGTH: 415
37 TYPE: PRT
38 ORGANISM: Drosophila
39 SEQUENCE 415 AA: 4700 MW: 87648.0N.

Query Match
  Best Local Similarity: 95.7% Score 45: DB 2: Length 216:
  Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

14 199 VRRPP 214
15 1 VVRRPP 6

RESULT 9
16 US-08-630-672A-2 STANDARD: PRT: 542 AA.
17 AC XXXXX
18 DT
19 DE
20 XX
21 XX
22 XX
23 SEQUENCE 2, Application US/08630672A
24 SEQUENCE 2, Application US/08630672A
25 GENERAL INFORMATION:
26 APPLICANT: Venter, J. Craig
27 APPLICANT: Vasilakos, Eric A.
28 APPLICANT: Plovy, Scott E.
29 APPLICANT: Yoon, Hyeon
30 APPLICANT: Rhee, Seung H.
31 TITLE OF INVENTION: Methods of Diagnosing a Clinical State
32 TITLE OF INVENTION: of Crohn's Disease with Patterns of Nucleotide Nucleic
33 NUMBER OF SEQUENCES: 12
34 CURRENT APPLICATION NUMBER:
35 ADDRESSEE: Cephalon and Fluor
36 STREET: 4370 La Jolla Village Drive, Suite 700
37 CITY: San Diego
38 STATE: California
39 COUNTRY: USA
40 ZIP: 92122
41 COMPUTER RECAPTURE FROM:
42 MEDIUM TYPE: Floppy disk

```

```

11 TYPE: PRT
12 ORGANISM: Drosophila
13 SEQUENCE 415 AA: 2354 MW: 24137.0N.

Query Match
  Best Local Similarity: 95.7% Score 45: DB 2: Length 216:
  Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

14 199 VRRPP 214
15 1 VVRRPP 6

RESULT 8
16 US-60-161-942-164 STANDARD: PRT: 415 AA.
17 AC XXXXX
18 DT
19 DE
20 XX
21 XX
22 XX
23 SEQUENCE 1964, Application US/60161942
24 SEQUENCE 1964, Application US/60161942
25 GENERAL INFORMATION:
26 APPLICANT: Venter, J. Craig
27 TITLE OF INVENTION: The Drosophila Genome, Primary Nucleic
28 TITLE OF INVENTION: Acid Sequences, Systems Containing the Nucleic Acid Sequen
29 TITLE OF INVENTION: and Uses Thereof
30 FILE REFERENCE: 1600122
31 CURRENT APPLICATION NUMBER: US/60161942
32 CURRENT FILING DATE: 1999-10-28
33 NUMBER OF SEQ ID NOS: 2620
34 SOFTWARE: FASTSEQ for Windows Version 4.0
35 SEQ ID NO: 1964
36 LENGTH: 415
37 TYPE: PRT
38 ORGANISM: Drosophila
39 SEQUENCE 415 AA: 4700 MW: 87648.0N.

Query Match
  Best Local Similarity: 95.7% Score 45: DB 2: Length 216:
  Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

14 199 VRRPP 214
15 1 VVRRPP 6

RESULT 9
16 US-08-630-672A-2 STANDARD: PRT: 542 AA.
17 AC XXXXX
18 DT
19 DE
20 XX
21 XX
22 XX
23 SEQUENCE 2, Application US/08630672A
24 SEQUENCE 2, Application US/08630672A
25 GENERAL INFORMATION:
26 APPLICANT: Venter, J. Craig
27 APPLICANT: Vasilakos, Eric A.
28 APPLICANT: Plovy, Scott E.
29 APPLICANT: Yoon, Hyeon
30 APPLICANT: Rhee, Seung H.
31 TITLE OF INVENTION: Methods of Diagnosing a Clinical State
32 TITLE OF INVENTION: of Crohn's Disease with Patterns of Nucleotide Nucleic
33 NUMBER OF SEQUENCES: 12
34 CURRENT APPLICATION NUMBER:
35 ADDRESSEE: Cephalon and Fluor
36 STREET: 4370 La Jolla Village Drive, Suite 700
37 CITY: San Diego
38 STATE: California
39 COUNTRY: USA
40 ZIP: 92122
41 COMPUTER RECAPTURE FROM:
42 MEDIUM TYPE: Floppy disk

```

CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patent In Release #1 0, Version #1 25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/99/000,572A  
 CC FILING DATE: 12-APR-1996

CC CLASSIFICATION: 435  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Campbell, Cathryn A.  
 CC REGISTRATION NUMBER: 31,815  
 CC REFERENCE/DOCKET NUMBER: P-PM 1997

CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (619) 535-9001  
 CC TELEFAX: (619) 535-8949

CC INFORMATION FOR SEQ ID NO: 2:  
 CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 532 amino acids

CC TYPE: amino acid

CC TOPOLOGY: linear

CC MOLECULE TYPE: Protein

CC SEQUENCE 532 AA: 58263 MW: 1541840 CN:

Query Match 95.7% Score 45. DB 11: length 532:  
 Best Local Similarity 83.3% Pred. No. 3329-02:  
 Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

Db 36 VIRPPP 41

UY 1 VIRPPP 6

RESULT 10

ID US-09-450-060-3924 STANDARD: PRI: 198 AA:

XX XXXXX

AC XXXXX

DT

XX

DE Sequence 3924, Application US/99/450969

XX Sequence 3924, Application US/99/450969

CC GENERAL INFORMATION:

CC APPLICANT: Lynn Doucette Stamm et al.

CC TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO SIAPHYLOCOCCUS

CC TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

CC FILE REFERENCE: PATH99-09A

CC CURRENT APPLICATION NUMBER: US/99/450 969

CC CURRENT FILING DATE: 1999-11-29

CC NUMBER OF SEQ ID NOS: 7544

CC SEQ ID NO 3924

LENGTH: 198

TYPE: PRI

CC ORGANISM: S.epidermidis

CC SEQUENCE 198 AA: 21600 MW: 222826 CN:

Query Match 91.5% Score 43: DB 19: Length 198:

Best Local Similarity 66.7% Pred. No. 5879-02:

Matches 4: Conservative 2: Mismatches 0: Indels 0: Gaps 0:

Db 127 VIRPPP 132

UY 1 VIRPPP 6

Search completed: Fri Feb 25 13:34:12 2000

Job time : 2.7 secs.





















1. The first part of the document discusses the importance of maintaining accurate records of all transactions and activities. It emphasizes that proper record-keeping is essential for transparency and accountability, particularly in financial matters. The text outlines various methods for organizing and storing data, including digital databases and physical filing systems. It also mentions the need for regular audits and reviews to ensure the integrity of the information.

2. The second section focuses on the role of communication in achieving organizational goals. It highlights the importance of clear and concise communication, both internally and externally. The text provides examples of effective communication strategies, such as regular team meetings, open-door policies, and the use of various communication channels like email, phone, and face-to-face interactions. It also discusses the importance of listening and understanding the needs and concerns of all stakeholders.

3. The third part of the document addresses the challenges of managing a large and diverse workforce. It discusses the importance of providing ongoing training and development opportunities to ensure that employees have the skills and knowledge needed to perform their jobs effectively. The text also touches on the importance of creating a positive work environment that fosters collaboration and innovation. It mentions the need for flexible work arrangements and the importance of recognizing and rewarding employee achievements.

4. The final section discusses the importance of staying up-to-date with the latest trends and technologies in the industry. It emphasizes that organizations must be proactive in adopting new technologies and processes to remain competitive. The text provides examples of emerging technologies and their potential applications in various industries. It also discusses the importance of having a clear vision and strategy for the future, and the need for continuous learning and improvement.



\*\*\*\*\*  
 WIRETEL  
 \*\*\*\*\*

Release 3.1A John P. Collins, Biocomputing Research Unit,  
 University of Edinburgh, University of Edinburgh, J.K.  
 Distribution rights by Oxford Molecular Ltd

Master file: protein - protein database search, using Smith-Waterman algorithm  
 Run on: Fri Feb 25 14:08:53 2000; Master time 1.67 Seconds  
 Tabular output not generated. 31,869 Million cell updates/sec

File: US-09-132-799-1  
 Description: 1347 from US09132799.rpt  
 Format: Sequence  
 1 VWPB 4

Sequence: 1347  
 1347

Search: 134718 seqs, 14297625 residues

Post processing: Minimum March 94  
 Clustering: 1000 Summaries

Database: 1347 MB 2.5H\_2OMB 3.6\_2OMB 4.0C19\_2OMB 5.Backfiles1

Statistics: Mean 11.064; Variance 28.015; scale 0.427

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

# SUMMARY

Result	N	Score	Length	ID	Description	Pred. No.
1	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
2	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
3	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
4	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
5	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
6	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
7	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
8	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
9	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
10	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
11	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
12	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
13	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
14	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
15	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
16	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
17	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
18	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
19	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
20	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
21	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
22	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
23	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
24	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
25	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
26	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
27	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
28	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
29	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
30	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
31	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
32	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
33	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
34	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
35	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
36	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
37	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
38	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
39	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
40	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
41	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
42	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
43	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
44	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
45	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
46	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
47	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
48	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
49	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
50	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
51	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
52	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
53	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
54	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
55	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
56	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
57	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
58	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
59	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
60	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
61	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
62	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
63	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
64	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
65	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
66	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
67	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
68	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
69	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
70	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
71	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
72	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
73	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
74	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
75	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
76	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
77	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
78	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
79	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
80	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
81	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
82	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
83	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
84	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
85	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
86	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
87	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
88	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
89	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
90	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
91	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
92	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
93	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
94	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
95	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02
96	41	100.0	12	4	US-08-548- Sequence 32, Applicati	4.64e-02

[illegible]



CC APPLICATION NUMBER: US09/448,540  
 CC FILING DATE: 15-SEP-1994  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 08/290,641  
 CC FILING DATE: 15-AUG-1994  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 07,779,223  
 CC FILING DATE: 15-OCT-1992  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Smith, William M.  
 CC REGISTRATION NUMBER: 30,223  
 CC REFERENCE/DOCKET NUMBER: 11509-50-1  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 415-326-2400  
 CC TELEFAX: 415-326-2422  
 CC INFORMATION FOR SEQ ID NO: 1:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 12 amino acids  
 CC TYPE: amino acid  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: peptide  
 CC IMMEDIATE SOURCE:  
 CC CLONING: 11509-50-1  
 CC SEQUENCE: 12 AA: 1376 MW: 912 CN:  
 CC  
 CC Query Match: 100.0% Score 31: PB 1: Length 12:  
 CC Best Local Similarity 100.0% Pred No: 4,640-02:  
 CC Matches: 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:  
 CC  
 Db 8 VWRP 11  
 QY 1 VWRP 4  
 CC  
 CC RESULT 3  
 CC ID: US-09-132-799-132 STANDARD: FBI: 12 AA:  
 CC AC xxxxxx  
 CC  
 CC Sequence 13: Application US/09779223  
 CC Patent No. 5070170  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Schatz, Peter J.  
 CC APPLICANT: Cull, Millard G.  
 CC APPLICANT: Miller, Jeff F.  
 CC TITLE OF INVENTION: Peptide Library and Screening Method  
 CC NUMBER OF SEQUENCES: 9  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: William M. Smith  
 CC STREET: One Market Plaza, Ste 201, Los Angeles 90001  
 CC CITY: San Francisco  
 CC STATE: California  
 CC COUNTRY: USA  
 CC ZIP: 94105  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patent Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA: 15-SEP-1994  
 CC APPLICANT NUMBER: 11509-50-1  
 CC FILING DATE: 15-SEP-1994  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 07,779,223  
 CC FILING DATE: 15-OCT-1991  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Smith, William M.  
 CC REGISTRATION NUMBER: 30,223  
 CC REFERENCE/DOCKET NUMBER: 11509-50-1  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 415-326-2400  
 CC TELEFAX: 415-326-2422  
 CC INFORMATION FOR SEQ ID NO: 1:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 12 amino acids  
 CC TYPE: AMINO ACID  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: linear

CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 415-326-2400  
 CC TELEFAX: 415-326-2422  
 CC INFORMATION FOR SEQ ID NO: 1:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 12 amino acids  
 CC TYPE: AMINO ACID  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: peptide  
 CC IMMEDIATE SOURCE:  
 CC CLONING: 11509-50-1  
 CC SEQUENCE: 12 AA: 1376 MW: 912 CN:  
 CC  
 CC Query Match: 100.0% Score 31: PB 1: Length 12:  
 CC Best Local Similarity 100.0% Pred No: 4,640-02:  
 CC Matches: 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:  
 CC  
 Db 8 VWRP 11  
 QY 1 VWRP 4  
 CC  
 CC RESULT 4  
 CC ID: US-09-132-799-132 STANDARD: FBI: 12 AA:  
 CC AC xxxxxx  
 CC  
 CC Sequence 12: Application US/07963321  
 CC Patent No. 5338665  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Schatz, Peter J.  
 CC APPLICANT: Cull, Millard G.  
 CC APPLICANT: Miller, Jeff F.  
 CC APPLICANT: Stemmer, Willem P.C.  
 CC TITLE OF INVENTION: Peptide Library and Screening Method  
 CC NUMBER OF SEQUENCES: 9  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: William M. Smith  
 CC STREET: One Market Plaza, Ste 201, Los Angeles 90001  
 CC CITY: San Francisco  
 CC STATE: California  
 CC COUNTRY: USA  
 CC ZIP: 94105  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patent Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA: 15-SEP-1994  
 CC APPLICANT NUMBER: 11509-50-1  
 CC FILING DATE: 15-SEP-1994  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 07,779,223  
 CC FILING DATE: 15-OCT-1991  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Smith, William M.  
 CC REGISTRATION NUMBER: 30,223  
 CC REFERENCE/DOCKET NUMBER: 11509-50-1  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 415-326-2400  
 CC TELEFAX: 415-326-2422  
 CC INFORMATION FOR SEQ ID NO: 1:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 12 amino acids  
 CC TYPE: AMINO ACID  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: linear

```

CC MOLECULE TYPE: Peptide
CC IMMEDIATE SOURCE:
CC SEQ ID: 1
CC SEQUENCE: 16 AA: 106 MW: 912.08
CC
CC Query Match: 100.0% Score 31: DB 1 Length 12:
CC Best Local Similarity 100.0% Pred. No. 4,646-42
CC Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0
CC
DB 1 VWRP 11
CC 1 VWRP 4
CC
CC RESULT 6
CC ID US-08-318-200-8 STANDARD: PRI: 16 AA.
CC XX
CC AC
CC XX
CC XX
CC DT
CC XX
CC DE
CC SEQUENCE 8: Application US/9804200
CC
CC Sequence 8: Application US/9804200
CC Patent No. 557499
CC GENERAL INFORMATION:
CC APPLICANT: Astris, M. Z.
CC APPLICANT: Ashizawa, T.
CC TITLE OF INVENTION: N-TERMINAL PVP-PEPTIDE COMPOUND
CC NUMBER OF INVENTION: 1
CC NUMBER OF SEQUENCES: 25
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Arnold, White & Dore
CC STREET: P.O. Box 443
CC CITY: Houston
CC STATE: Texas
CC COUNTRY: USA
CC ZIP: 77210
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: 09/87779-223
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 09/799-132
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Michael, C. Stover
CC TELEPHONE: 713-787-1400
CC TELEFAX: 713-789-2479
CC INFORMATION P.F. SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 16 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC SEQUENCE 16 AA: 106 MW: 1478.08
CC
CC Query Match: 100.0% Score 31: DB 1 Length 12:
CC Best Local Similarity 100.0% Pred. No. 4,646-42
CC Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0
CC
DB 7 VWRP 10
CC 1 VWRP 4
CC
CC RESULT 7
CC ID US-09-047-125-14 STANDARD: PRI: 92 AA.
CC XX
CC AC
CC XX
CC DT
CC XX
CC DE
CC SEQUENCE 14: Application US/9847125

```

XX Sequence 14, Application US/08/52816A  
 CC Patent No. 5872215  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Leand F. Vellera, Peter Brimovskis,  
 CC TITLE OF INVENTION: Mark's Disease Repressors  
 CC NUMBER OF SEQUENCES: 4  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Ian C. McLeod  
 CC STREET: 2190 Commons Parkway  
 CC CITY: Okemos  
 CC STATE: Michigan  
 CC COUNTRY: USA  
 CC ZIP: 48864  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Diskette, 5.25 inch, 1.2 MB  
 CC FORM TYPE: IBM PC, MS-DOS 5.0  
 CC OPERATING SYSTEM: MS-DOS 5.0  
 CC SOFTWARE: PC-Write 3.02  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/09/132-799  
 CC FILING DATE: 23-MAY-1996  
 CC CLASSIFICATION:  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: 07/070,425  
 CC FILING DATE: 23-MAY-1991  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Ian C. McLeod  
 CC REGISTRATION NUMBER: 26,991  
 CC REFERENCE/DOCKET NUMBER: MSU 4,1-132  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (517) 347-4100  
 CC TELEFAX: (517) 347-4103  
 CC INFORMATION FOR SEQ ID NO: 14:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 92 amino acids  
 CC TYPE: amino acid  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: peptide  
 CC FRAGMENT TYPE: N-terminal fragment  
 CC ORIGINAL SOURCE:  
 CC ORGANISM: herpes simplex virus type 1 (HSV1)  
 CC FEATURE:  
 CC NAME, FLAG, Position of HSV1 gene US1 flanked by  
 CC LOCATION: 1-16-92  
 CC OTHER INFORMATION: Peptide homologous to the US2 gene  
 CC OTHER INFORMATION: Peptide homologous to the US2 gene

SO SEQUENCE 92 AA, 10121 MW, 43655 GR.  
 Query Match 100.0%, Score 31, DB 2, Length 92;  
 Best Local Similarity 100.0%, Pred. No. 4 64e-02;  
 Matches 4: Conservative 0, Mismatches 0, Indels 0, Gaps 0.

Db 50 VWRP 53  
 QY 1 VWRP 4  
 RESULT 8  
 ID US-08-552-815A-14 STANDARD: PRT: 116 AA.  
 XX XXXXXX  
 XX XXXXXX

XX Sequence 14, Application US/08/52816A  
 CC Patent No. 5872215  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Leand F. Vellera, Peter Brimovskis,  
 CC TITLE OF INVENTION: Mark's Disease Repressors  
 CC NUMBER OF SEQUENCES: 4  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Ian C. McLeod  
 CC STREET: 2190 Commons Parkway  
 CC CITY: Okemos  
 CC STATE: Michigan  
 CC COUNTRY: USA  
 CC ZIP: 48864

CC APPLICANT: McCaifferty, JG  
 CC TITLE OF INVENTION: Protein Binding Methods - Anticatalytic and  
 CC TITLE OF INVENTION: Methods  
 CC NUMBER OF SEQUENCES: 53  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Mark's Disease Repressors  
 CC STREET: 4500 Sears Road, 2nd South Market Drive  
 CC CITY: Chicago  
 CC STATE: Illinois  
 CC COUNTRY: United States of America  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC OPERATING SYSTEM: MS-DOS/MS-WIN  
 CC SOFTWARE: Patent In Release #1.0, Version #1.25 (RPO)  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/09/132-799  
 CC FILING DATE: 23-MAY-1996  
 CC CLASSIFICATION:  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: 07/070,425  
 CC FILING DATE: 23-MAY-1991  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: David W. Clough  
 CC REGISTRATION NUMBER: 36,107  
 CC REFERENCE/DOCKET NUMBER: 2011/33308  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (312) 424-0300  
 CC INFORMATION FOR SEQ ID NO: 14:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 116 amino acids  
 CC TYPE: amino acid  
 CC STRANDEDNESS: linear  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: linear  
 CC FRAGMENT TYPE: N-terminal fragment  
 CC ORIGINAL SOURCE:  
 CC ORGANISM: herpes simplex virus type 1 (HSV1)  
 CC FEATURE:  
 CC NAME, FLAG, Position of HSV1 gene US1 flanked by  
 CC LOCATION: 1-16-92  
 CC OTHER INFORMATION: Peptide homologous to the US2 gene  
 CC OTHER INFORMATION: Peptide homologous to the US2 gene

SO SEQUENCE 116 AA, 12860 MW, 59765 GR.  
 Query Match 100.0%, Score 31, DB 2, Length 116;  
 Best Local Similarity 100.0%, Pred. No. 4 64e-02;  
 Matches 4: Conservative 0, Mismatches 0, Indels 0, Gaps 0.

Db 11 VWRP 14  
 QY 1 VWRP 4  
 RESULT 9  
 ID US-08-828-832-3 STANDARD: PRT: 158 AA.  
 XX XXXXXX  
 XX XXXXXX

XX Sequence 14, Application US/08/52816A  
 CC Patent No. 5872215  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Leand F. Vellera, Peter Brimovskis,  
 CC TITLE OF INVENTION: Mark's Disease Repressors  
 CC NUMBER OF SEQUENCES: 4  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Ian C. McLeod  
 CC STREET: 2190 Commons Parkway  
 CC CITY: Okemos  
 CC STATE: Michigan  
 CC COUNTRY: USA  
 CC ZIP: 48864







\*\* RESUME SCIENTIFIC ADDRESS:  
 \*\* ADDRESSER: FREDERICK W. FETTER, JR.  
 \*\* STREET: 11545 W. BERNARDO COURT, STE. 302  
 \*\* CITY: SAN DIEGO  
 \*\* STATE: CA  
 \*\* ZIP: 92127  
 \*\* COUNTRY: USA  
 \*\* PHONE: 619-441-2700  
 \*\* FAX: 619-441-2700  
 \*\* E-MAIL: FETTER@FETTER.COM  
 \*\* OPERATING SYSTEM: PC/DOS/MS-DOS  
 \*\* SOFTWARE: Patent In Release #1.0, Version #1.25  
 \*\* CURRENT APPLICATION DATA:  
 \*\* APPLICATION NUMBER: 92/0224.01  
 \*\* FILING DATE: 19921013  
 \*\* CLASSIFICATION: 4.5  
 \*\* ATTORNEY/AGENT INFORMATION:  
 \*\* NAME: PETER H.D. FREDERICK W.  
 \*\* REGISTRATION NUMBER: 31,286  
 \*\* REFERENCE NUMBER: 92/0224.01  
 \*\* TITLE: INFORMATION INFORMATION  
 \*\* TELEPHONE: (619) 451-1120  
 \*\* TELEFAX: (619) 451-9628  
 \*\* INFORMATION FOR SEQ ID NO: 1:  
 \*\* SEQUENCE CHARACTERISTICS:  
 \*\* LENGTH: 192 amino acids  
 \*\* TYPE: AMINO ACID  
 \*\* TOPOLOGY: Linear  
 \*\* MOLECULE TYPE: Protein  
 \*\* DESCRIPTION:  
 \*\* HYPOTHETICAL:  
 \*\* ANTI-SENSE:  
 \*\* FRAGMENT TYPE:  
 \*\* ORIGINAL SOURCE:  
 \*\* ORGANISM:  
 \*\* STRAIN:  
 \*\* INDIVIDUAL ISOLATE:  
 \*\* DEVELOPMENTAL STAGE:  
 \*\* HAPLOTYPE:  
 \*\* TISSUE TYPE:  
 \*\* CELL TYPE:  
 \*\* CELL LINE:  
 \*\* ORGANELLER:  
 \*\* IMMEDIATE SOURCE:  
 \*\* CLONE:  
 \*\* POSITION IN SEQUENCE:  
 \*\* CHROMOSOME/SEGMENT:  
 \*\* MAP POSITION:  
 \*\* UNITS:  
 \*\* FEATURE:  
 \*\* NAME/KEY:  
 \*\* LOCATION:  
 \*\* IDENTIFICATION METHOD:  
 \*\* OTHER INFORMATION:  
 \*\* PUBLICATION INFORMATION:  
 \*\* AUTHOR: L. Hansson et al  
 \*\* TITLE: DNA Encoding Kappa Cassin. Process for obtaining the protein and

RESULT 14  
 Query Match 100.0% Score 31; DB 1; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 4,640-02;  
 Matches 4; Conservation 1; Mismatches 0; Gaps 0.

DB 86 VWRP 89  
 QY 1 VWRP 4

RESULT 15  
 Query Match 100.0% Score 31; DB 1; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 4,640-02;  
 Matches 4; Conservation 1; Mismatches 0; Gaps 0.

DB 86 VWRP 89  
 QY 1 VWRP 4

RESULT 16  
 Query Match 100.0% Score 31; DB 1; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 4,640-02;  
 Matches 4; Conservation 1; Mismatches 0; Gaps 0.

DB 86 VWRP 89  
 QY 1 VWRP 4

RESULT 17  
 Query Match 100.0% Score 31; DB 1; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 4,640-02;  
 Matches 4; Conservation 1; Mismatches 0; Gaps 0.

DB 86 VWRP 89  
 QY 1 VWRP 4

RESULT 18  
 Query Match 100.0% Score 31; DB 1; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 4,640-02;  
 Matches 4; Conservation 1; Mismatches 0; Gaps 0.

DB 86 VWRP 89  
 QY 1 VWRP 4

RESULT 19  
 Query Match 100.0% Score 31; DB 1; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 4,640-02;  
 Matches 4; Conservation 1; Mismatches 0; Gaps 0.

DB 86 VWRP 89  
 QY 1 VWRP 4

RESULT 20  
 Query Match 100.0% Score 31; DB 1; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 4,640-02;  
 Matches 4; Conservation 1; Mismatches 0; Gaps 0.

DB 86 VWRP 89  
 QY 1 VWRP 4

RESULT 21  
 Query Match 100.0% Score 31; DB 1; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 4,640-02;  
 Matches 4; Conservation 1; Mismatches 0; Gaps 0.

DB 86 VWRP 89  
 QY 1 VWRP 4

RESULT 22  
 Query Match 100.0% Score 31; DB 1; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 4,640-02;  
 Matches 4; Conservation 1; Mismatches 0; Gaps 0.

DB 86 VWRP 89  
 QY 1 VWRP 4

\*\* RESUME SCIENTIFIC ADDRESS:  
 \*\* ADDRESSER: FREDERICK W. FETTER, JR.  
 \*\* STREET: 11545 W. BERNARDO COURT, STE. 302  
 \*\* CITY: SAN DIEGO  
 \*\* STATE: CA  
 \*\* ZIP: 92127  
 \*\* COUNTRY: USA  
 \*\* PHONE: 619-441-2700  
 \*\* FAX: 619-441-2700  
 \*\* E-MAIL: FETTER@FETTER.COM  
 \*\* OPERATING SYSTEM: PC/DOS/MS-DOS  
 \*\* SOFTWARE: Patent In Release #1.0, Version #1.25  
 \*\* CURRENT APPLICATION DATA:  
 \*\* APPLICATION NUMBER: 92/0224.01  
 \*\* FILING DATE: 19921013  
 \*\* CLASSIFICATION: 4.5  
 \*\* ATTORNEY/AGENT INFORMATION:  
 \*\* NAME: PETER H.D. FREDERICK W.  
 \*\* REGISTRATION NUMBER: 31,286  
 \*\* REFERENCE NUMBER: 92/0224.01  
 \*\* TITLE: INFORMATION INFORMATION  
 \*\* TELEPHONE: (619) 451-1120  
 \*\* TELEFAX: (619) 451-9628  
 \*\* INFORMATION FOR SEQ ID NO: 1:  
 \*\* SEQUENCE CHARACTERISTICS:  
 \*\* LENGTH: 192 amino acids  
 \*\* TYPE: AMINO ACID  
 \*\* TOPOLOGY: Linear  
 \*\* MOLECULE TYPE: Protein  
 \*\* DESCRIPTION:  
 \*\* HYPOTHETICAL:  
 \*\* ANTI-SENSE:  
 \*\* FRAGMENT TYPE:  
 \*\* ORIGINAL SOURCE:  
 \*\* ORGANISM:  
 \*\* STRAIN:  
 \*\* INDIVIDUAL ISOLATE:  
 \*\* DEVELOPMENTAL STAGE:  
 \*\* HAPLOTYPE:  
 \*\* TISSUE TYPE:  
 \*\* CELL TYPE:  
 \*\* CELL LINE:  
 \*\* ORGANELLER:  
 \*\* IMMEDIATE SOURCE:  
 \*\* CLONE:  
 \*\* POSITION IN SEQUENCE:  
 \*\* CHROMOSOME/SEGMENT:  
 \*\* MAP POSITION:  
 \*\* UNITS:  
 \*\* FEATURE:  
 \*\* NAME/KEY:  
 \*\* LOCATION:  
 \*\* IDENTIFICATION METHOD:  
 \*\* OTHER INFORMATION:  
 \*\* PUBLICATION INFORMATION:  
 \*\* AUTHOR: L. Hansson et al  
 \*\* TITLE: DNA Encoding Kappa Cassin. Process for obtaining the protein and

Sequence 1, Application: US/09-132-799-1

```

XX Sequence 1: Application US/08865336
CC Patent No. 5958725
CC GENERAL INFORMATION:
CC APPLICANT: Bandman, Olga
CC APPLICANT: Guedler, Karl J.
CC APPLICANT: Shah, Purvi
CC APPLICANT: Pettithory, Joanne R.
CC APPLICANT: Corley, Neil C.
CC TITLE OF INVENTION: NEW HUMAN Dp1 HOMOLOG
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Insyle Pharmaceuticals, Inc.
CC STREET: 3174 Porter Drive
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: PASTSEQ for Windows Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/485,336
CC FILING DATE: Herewith
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC AGENCY/AGENT INFORMATION:
CC NAME: Billings, Lucy J.
CC REGISTRATION NUMBER: 36,749
CC REFERENCE/BACKL NUMBER: PF-0313 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-855-0555
CC TELEFAX: 415-845-4166
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 184 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC LIBRARY: SINING102
CC CLONE: 236773
SQ SEQUENCE 184 AA: 20733 MW: 175484 CN:

Query Match 100.0% Score 13189.27 Length 184
Best Local Similarity 100.0% Pred No. 4.64e+02
Matches 4: Constitutive 0: Mismatches 0: Indels 0: Gaps 0;

Db 139 VWRP 142
QY 1 VWRP 4

```

Search completed: Fri Feb 25 13:09:37 2000  
Job time: 44 secs.

# WISWIRE

(1X)

Release 3.1A John F. Collins, Bioinformatics Research Unit,

Copyright (C) 1994-1998 University of Edinburgh, U.K.

Distribution rights by Oxford Molecular Ltd

WISWIRE is a protein database search, using Smith-Waterman algorithm

Run on: Fri Feb 25 14:05:17 2000; Maspar time 2.97 Seconds

Library output not generated. 53,879 Million cell updates/sec

Title: 48599-142-799-1

Description: (1-4) from US09132799.ppt

Project Name: N

Sequence: 1 VVWP 4

Search table: FAM 160

Cap 16

Search: 122910 seqs, 4096804 residues

Post processing: Minimum March 90X

MS-09-132-799-1

Database: PIR

Query: 13p11.2ip12.2ip13.4ip14

Statistics: Mean 17.13, Variance 15.436, S.d. 1.244

Pred. No. in the number of results predicted by above to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	N	Score	Match	Length	IR	ID	Description	Pred. No.
1	41	120.0	16	3	PH1778		r cell receptor alpha	2,536-02
2	41	100.0	20	4	PH1793		r cell receptor alpha	2,536-02
3	41	100.0	24	2	A27242		nicotinic acetylcholine	2,536-02
4	41	100.0	24	2	A27242		nicotinic acetylcholine	2,536-02
5	41	100.0	66	2	S47377		hypothetical protein	2,536-02
6	41	100.0	69	2	J04545		hypothetical protein	2,536-02
7	41	100.0	72	2	A41884		positive regulator of	2,536-02
8	41	100.0	74	2	G71771		probable transmembrane	2,536-02
9	41	100.0	82	2	S75451		hypothetical protein	2,536-02
10	41	100.0	87	2	G70782		hypothetical protein	2,536-02
11	41	100.0	96	2	H49042		ribosomal protein L21	2,536-02
12	41	100.0	98	2	S26928		lg heavy chain V red1	2,536-02
13	41	100.0	103	2	S16954		hypothetical protein	2,536-02
14	41	100.0	103	2	S22563		hypothetical protein	2,536-02
15	41	100.0	103	2	P73897		hypothetical protein	2,536-02
16	41	100.0	106	2	G64111		probable transmembrane	2,536-02
17	41	100.0	112	2	PH1654		lg heavy chain V red1	2,536-02
18	41	100.0	122	2	P76510		hypothetical protein	2,536-02
19	41	100.0	122	2	P76510		hypothetical protein	2,536-02
20	41	100.0	123	2	A31172		trans protein - Escher	2,536-02
21	41	100.0	126	2	S41327		gene H2 protein - Wye	2,536-02
22	41	100.0	142	2	I47143		aspartic protein precursor	2,536-02
23	41	100.0	144	2	G73657		hypothetical protein	2,536-02

24	31	100.0	146	1	HVWSB1	in heavy chain protein	2,536-02
25	31	100.0	146	2	IN0824	tetracycline-inducible	2,536-02
26	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
27	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
28	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
29	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
30	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
31	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
32	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
33	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
34	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
35	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
36	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
37	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
38	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
39	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
40	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
41	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
42	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
43	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
44	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
45	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
46	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
47	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
48	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
49	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
50	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
51	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
52	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
53	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
54	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
55	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
56	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
57	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
58	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
59	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
60	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
61	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
62	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
63	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
64	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
65	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
66	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
67	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
68	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
69	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
70	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
71	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
72	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
73	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
74	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
75	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
76	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
77	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
78	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
79	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
80	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
81	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
82	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
83	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
84	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
85	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
86	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
87	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
88	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
89	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
90	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
91	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
92	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
93	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
94	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
95	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02
96	31	100.0	144	2	S71506	ribosomal protein L16	2,536-02





[illegible]



[illegible]





## SUMMARY      #length 24    #checksum 2911

Query Match      100.0%    Score 21    DB 2    Length 24.  
 Best Local Similarity 100.0%    Pref No. 2530-02  
 Matches      4:    Conservative      0:    Mismatches      0:    Indels      0:    Gaps      0:

Db    18 VWRP 21  
 ||||  
 QY    1 VWRP 4

RESULT      4      A25738      #type fragment  
 ENTRY      nicotinic acetylcholine receptor alpha chain - chicken  
 TITLE      (fragment)  
 ORGANISM      #molecule\_type DNA  
 #author      #formal\_name Gallus gallus #checksum 2911  
 #journal      20-Jul-1987 #sequence\_revision 21-May-1988 #text\_change  
 #title      15-Jun-1996  
 #cross-references MIM:6020494  
 #accession      A25738  
 #molecule\_type protein  
 #residues      124 ##label CON

COMMENT      The source of this protein was muscle.  
 CLASSIFICATION      #superfamily acetylcholine receptor  
 KEYWORDS      glycoprotein, ion channel, neurotransmitter receptor,  
 postsynaptic membrane, transmembrane protein  
 SUMMARY      #length 24    #checksum 3483

Query Match      100.0%    Score 31    DB 2    Length 24.  
 Best Local Similarity 100.0%    Pref No. 2530-02  
 Matches      4:    Conservative      0:    Mismatches      0:    Indels      0:    Gaps      0:

Db    17 VWRP 20  
 ||||  
 QY    1 VWRP 4

RESULT      1      S45377      #type complete  
 ENTRY      hypothetical protein Acm16 - Anopheles stephensi  
 TITLE      #formal\_name Anopheles stephensi  
 ORGANISM      16-Mar-1999 #sequence\_revision 17-Mar-1999 #text\_change  
 #journal      Samant, S.A.; Vilekar, N.R.; Lole, K.S.; Deobagkar, D.D.;  
 #title      Deobagkar, D.N.;  
 #cross-references MIM:6020494  
 #accession      S45377  
 #molecule\_type DNA  
 #residues      1-66 ##label SAM  
 #cross-references DML:227581, MIM:6020494, 17-Mar-1999  
 #experimental\_source strain liston  
 #length 66    #molecular\_weight 7950    #checksum 3020

Query Match      100.0%    Score 31    DB 2    Length 66:  
 Best Local Similarity 100.0%    Pref No. 2530-02  
 Matches      4:    Conservative      0:    Mismatches      0:    Indels      0:    Gaps      0:

Db    17 VWRP 20  
 ||||  
 QY    1 VWRP 4

RESULT      1      S45377      #type complete  
 ENTRY      hypothetical protein Acm16 - Anopheles stephensi  
 TITLE      #formal\_name Anopheles stephensi  
 ORGANISM      16-Mar-1999 #sequence\_revision 17-Mar-1999 #text\_change  
 #journal      Samant, S.A.; Vilekar, N.R.; Lole, K.S.; Deobagkar, D.D.;  
 #title      Deobagkar, D.N.;  
 #cross-references MIM:6020494  
 #accession      S45377  
 #molecule\_type DNA  
 #residues      1-66 ##label SAM  
 #cross-references DML:227581, MIM:6020494, 17-Mar-1999  
 #experimental\_source strain liston  
 #length 66    #molecular\_weight 7950    #checksum 3020

RESULT      6      J04545      #type complete  
 ENTRY      hypothetical 7 ok protein - Barophilic bacterium  
 TITLE      pressure-regulated operon  
 ORGANISM      #formal\_name Barophilic bacterium  
 DATE      29-Feb-1996 #sequence\_revision 19-Apr-1996 #text\_change  
 31-Oct-1997

ACCESSIONS      J04545  
 REFERENCE      J04545  
 #author      Kato, C.; Smorawinska, M.; Sato, T.; Horikoshi, K.  
 #journal      Biosci. Biotechnol. Biochem. (1996) 60:166-168  
 #title      Analysis of a pressure-regulated operon from the barophilic  
 bacterium strain DB6705.  
 #cross-references MIM:6020494  
 #accession      J04545  
 #molecule\_type DNA  
 #residues      1-69 ##label KAT  
 #cross-references DB:66756, EMBL:137777, MIM:6020494, 19-Apr-1996, 31-Oct-1997

SUMMARY      #experimental\_source DB6705  
 #length 69    #molecular\_weight 7001    #checksum 1843

Query Match      100.0%    Score 31    DB 2    Length 69:  
 Best Local Similarity 100.0%    Pref No. 2530-02  
 Matches      4:    Conservative      0:    Mismatches      0:    Indels      0:    Gaps      0:

Db    35 VWRP 38  
 ||||  
 QY    1 VWRP 4

RESULT      7      A41889      #type complete  
 ENTRY      positive regulator of phase P2 late gene transcription -  
 TITLE      Escherichia coli  
 ALTERNATE\_NAMES      rgt protein  
 ORGANISM      #formal\_name Escherichia coli  
 DATE      30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change  
 17-Mar-1999

ACCESSIONS      A41889; A41890; A54975  
 REFERENCE      A41889  
 #author      Barreiro, V.; Haggard-Ljungquist, E.  
 #journal      J. Bacteriol. (1992) 174:4086-4093  
 #title      Attachment sites for bacteriophage P2 on the Escherichia coli  
 chromosome: DNA sequences, localization on the physical  
 map, and detection of a P2-like remnant in Escherichia coli  
 K-12 derivatives.  
 #cross-references MIM:6020494  
 #accession      A41889  
 #status      preliminary; not compared with conceptual translation  
 #molecule\_type DNA  
 #residues      1-72 ##label RAP

REFERENCE      A41890  
 #author      Saito, A.; Ishida, Y.; Kato, T.; Barreiro, V.; Haggard-Ljungquist, E.  
 #journal      J. Bacteriol. (1992) 174:4094-4100  
 #title      Escherichia coli P2 attachment sites on the physical map  
 P2-odd genes.  
 #cross-references MIM:6020494  
 #accession      A41890  
 #status      preliminary  
 #molecule\_type DNA  
 #residues      1-72 ##label SLE

REFERENCE      A54975  
 #author      Barreiro, V.; Haggard-Ljungquist, E.; Kato, T.; Ishida, Y.; Saito, A.;  
 #journal      J. Bacteriol. (1992) 174:4086-4093  
 #title      Attachment sites for bacteriophage P2 on the Escherichia coli  
 chromosome: DNA sequences, localization on the physical  
 map, and detection of a P2-like remnant in Escherichia coli  
 K-12 derivatives.  
 #cross-references MIM:6020494  
 #accession      A41889

Query Match      100.0%    Score 31    DB 2    Length 69:  
 Best Local Similarity 100.0%    Pref No. 2530-02  
 Matches      4:    Conservative      0:    Mismatches      0:    Indels      0:    Gaps      0:

Db    35 VWRP 38  
 ||||  
 QY    1 VWRP 4

RESULT      7      A41889      #type complete  
 ENTRY      positive regulator of phase P2 late gene transcription -  
 TITLE      Escherichia coli  
 ALTERNATE\_NAMES      rgt protein  
 ORGANISM      #formal\_name Escherichia coli  
 DATE      30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change  
 17-Mar-1999

ACCESSIONS      A41889; A41890; A54975  
 REFERENCE      A41889  
 #author      Barreiro, V.; Haggard-Ljungquist, E.  
 #journal      J. Bacteriol. (1992) 174:4086-4093  
 #title      Attachment sites for bacteriophage P2 on the Escherichia coli  
 chromosome: DNA sequences, localization on the physical  
 map, and detection of a P2-like remnant in Escherichia coli  
 K-12 derivatives.  
 #cross-references MIM:6020494  
 #accession      A41889  
 #status      preliminary; not compared with conceptual translation  
 #molecule\_type DNA  
 #residues      1-72 ##label RAP

REFERENCE      A41890  
 #author      Saito, A.; Ishida, Y.; Kato, T.; Barreiro, V.; Haggard-Ljungquist, E.  
 #journal      J. Bacteriol. (1992) 174:4094-4100  
 #title      Escherichia coli P2 attachment sites on the physical map  
 P2-odd genes.  
 #cross-references MIM:6020494  
 #accession      A41890  
 #status      preliminary  
 #molecule\_type DNA  
 #residues      1-72 ##label SLE

REFERENCE      A54975  
 #author      Barreiro, V.; Haggard-Ljungquist, E.; Kato, T.; Ishida, Y.; Saito, A.;  
 #journal      J. Bacteriol. (1992) 174:4086-4093  
 #title      Attachment sites for bacteriophage P2 on the Escherichia coli  
 chromosome: DNA sequences, localization on the physical  
 map, and detection of a P2-like remnant in Escherichia coli  
 K-12 derivatives.  
 #cross-references MIM:6020494  
 #accession      A41889

Query Match      100.0%    Score 31    DB 2    Length 69:  
 Best Local Similarity 100.0%    Pref No. 2530-02  
 Matches      4:    Conservative      0:    Mismatches      0:    Indels      0:    Gaps      0:

Db    35 VWRP 38  
 ||||  
 QY    1 VWRP 4

[illegible]

```

Matches      4:  Conservative      0:  Mismatches      0:  Indels      0:  Gaps      0:

Db      75 VWRP 78
      ||||
QY      1 VWRP 4

RESULT      11
ENTRY      ribosomal protein l21 - Methanobacterium thermoautotrophicum
TITLE      (strain Delta H)
ORGANISM    #formal_name Methanobacterium thermoautotrophicum
DATE        05-Dec-1993 #sequence_revision 05 Feb-1997 #text_change
ACCESSIONS  H59042
REFERENCE    A59000
#authors    Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, G.; Lee, H.;
            Dubois, J.; Alaredge, T.; Bashir-Zadeh, R.; Blakes, D.;
            Cook, P.; Gilbert, W.; Harrison, D.; Hoang, I.; Keagle, P.;
            Lum, W.; Pothier, R.; Qiu, D.; Spadafora, F.; Vicaire, P.;
            Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwanji, N.; Caruso,
            A.; Bush, D.; Safer, H.; Patwell, D.; Prabhakar, S.;
            McLaughlin, J.; Shirai, S.; Goyal, A.; Piotrowski, S.;
            Church, G.W.; Cantale, G.T.; Moore, C.; Rice, F.; Neill,
            J.; Reeve, J.N.
#journal     J. Bacteriol. (1997) 179:7135-7155
#title       Complete genome sequence of Methanobacterium
            thermoautotrophicum Delta H: functional analysis and
            comparative genomics.
#cross-references MITH:98037514
#accession   H59042
#status      Preliminary; nucleic acid sequence not shown;
            translation not shown

#molecule_type DNA
#residues    1996 #label MTH
#cross-references EF A59000, CP A59000, NID 914810, FID 9252423
#experimental_source strain Delta H

GENETICS
#gene        MTH1323
#superfamily Halsecula ribosomal protein HL31
CLASSIFICATION
#length 96 #molecular-weight 11116 #checksum 6593

Query Match      100.0%  Score 31:  DB 2:  Length 96:
Best Local Similarity 100.0%  Prod No 258e-02:
Matches      4:  Conservative      0:  Mismatches      0:  Indels      0:  Gaps      0:

Db      86 VWRP 84
      |||
QY      1 VWRP 4

RESULT      12
ENTRY      lg heavy chain V region (DP-22) - human (fragment)
TITLE
ORGANISM    #formal_name Homo sapiens #common_name man
DATE        22-Nov-1993 #sequence_revision 10 Nov-1999 #text_change
ACCESSIONS  S26928
REFERENCE    S26885
#authors    Tomlinson, I.M.; Walter, G.; Marks, J.P.; Hleawlyn, M.R.;
            Winter, G.
#journal     J. Mol. Biol. (1997) 277:77-90
#title       The repertoire of human germline V(H) sequences reveals about
            fifty groups of V(H) segments with different hypervariable
            loops.
#cross-references MITH:98037514
#accession   S26928
#status      Preliminary
#molecule_type DNA
#residues    198 #label IOM
#cross-references EMBL 012934, NID 912934, FID 912934
CLASSIFICATION
#superfamily immunoglobulin V region; immunoglobulin homology
KEYWORDS    heterotetramer; immunoglobulin

```

```

FEATURE
15-99      #domain immunoglobulin homology #label IMM
SUMMARY    #length 96 #checksum 2406

Query Match      100.0%  Score 31:  DB 2:  Length 98:
Best Local Similarity 100.0%  Prod No 2.53e-02:
Matches      4:  Conservative      0:  Mismatches      0:  Indels      0:  Gaps      0:

Db      11 VWRP 14
      ||||
QY      1 VWRP 4

RESULT      13
ENTRY      S16054 #type complete
TITLE      Hypothetical protein 2 - phage phi-C31
ORGANISM    #formal_name Phage phi-C31
DATE        20 Feb-1999 #sequence_revision 20 Feb-1999 #text_change
ACCESSIONS  S16054
REFERENCE    S16053
#authors    Rausch, H.
#journal     Submitted to EMBL Data Library December 1999
#title       Submission
#accession   S16054
#status      Preliminary
#molecule_type DNA
#residues    1-100 #label RAU
#cross-references EMBL X57836, NID 914810, FID 9579069
GENETICS
#start_codon GTG
SUMMARY    #length 100 #molecular-weight 11548 #checksum 22

Query Match      100.0%  Score 31:  DB 2:  Length 100:
Best Local Similarity 100.0%  Prod No 2.53e-02:
Matches      4:  Conservative      0:  Mismatches      0:  Indels      0:  Gaps      0:

Db      30 VWRP 33
      ||||
QY      1 VWRP 4

RESULT      14
ENTRY      S22567 #type complete
TITLE      Hypothetical protein 2 - phage phi-C31
ORGANISM    #formal_name Phage phi-C31
DATE        22-Nov-1999 #sequence_revision 10-Nov-1999 #text_change
ACCESSIONS  S22567
REFERENCE    S22566
#authors    Rausch, H.; Lehmann, M.
#journal     Nucleic Acids Res (1991) 19:5187-5189
#title       Structural analysis of the actinophage phiC31 attachment
            site.
#cross-references MITH:92000003
#accession   S22567
#status      Preliminary; nucleic acid sequence not shown;
            translation not shown
#molecule_type DNA
#residues    1-100 #label RAU
#cross-references EMBL X57836
#note        the nucleotide sequence was submitted to the EMBL Data
            Library December 1990
SUMMARY    #length 100 #molecular-weight 11616 #checksum 41

Query Match      100.0%  Score 31:  DB 2:  Length 100:
Best Local Similarity 100.0%  Prod No 2.53e-02:
Matches      4:  Conservative      0:  Mismatches      0:  Indels      0:  Gaps      0:

Db      30 VWRP 33
      ||||
QY      1 VWRP 4

```





\*\*\*\*\*  
 WATSON  
 (19)  
 \*\*\*\*\*

Release 1.0A John F. Collins, Biocomputing Research Unit.

Copyright (c) 1998 University of Edinburgh, U.K.

Distribution rights by Oxford Molecular Ltd

WATSON: protein database search, using Smith-Waterman algorithm

Run on: Fri Feb 25 14:06:22 2000; MasPar time 2.60 Seconds  
 56.566 Million cell updates/sec

Protein source not recognized.

Protein: P08063 (14) from RS09142799.pep  
 Protein length: 31  
 Alignment: 1 WVF 4

Sequence: IAM 15  
 Gap 15

Scoring: 1000 steps, 2826243 residues

Scoring: Minimum Match 9%

Protein: P08063 (14) from RS09142799.pep  
 Protein length: 31  
 Alignment: 1 WVF 4

Statistics: Mean 18.116; Variance 13.185; scale 1.374

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Protein	No.	Score	Match	Length	PR	ID	Description	Pred. No.
1	41	100.0	24	31	100.0	1	ATM_HUMAN	1.03e-02
2	41	100.0	24	31	100.0	2	ATM_HUMAN	1.03e-02
3	41	100.0	24	31	100.0	3	ATM_HUMAN	1.03e-02
4	41	100.0	24	31	100.0	4	ATM_HUMAN	1.03e-02
5	41	100.0	24	31	100.0	5	ATM_HUMAN	1.03e-02
6	41	100.0	24	31	100.0	6	ATM_HUMAN	1.03e-02
7	41	100.0	24	31	100.0	7	ATM_HUMAN	1.03e-02
8	41	100.0	24	31	100.0	8	ATM_HUMAN	1.03e-02
9	41	100.0	24	31	100.0	9	ATM_HUMAN	1.03e-02
10	41	100.0	24	31	100.0	10	ATM_HUMAN	1.03e-02
11	41	100.0	24	31	100.0	11	ATM_HUMAN	1.03e-02
12	41	100.0	24	31	100.0	12	ATM_HUMAN	1.03e-02
13	41	100.0	24	31	100.0	13	ATM_HUMAN	1.03e-02
14	41	100.0	24	31	100.0	14	ATM_HUMAN	1.03e-02
15	41	100.0	24	31	100.0	15	ATM_HUMAN	1.03e-02
16	41	100.0	24	31	100.0	16	ATM_HUMAN	1.03e-02
17	41	100.0	24	31	100.0	17	ATM_HUMAN	1.03e-02
18	41	100.0	24	31	100.0	18	ATM_HUMAN	1.03e-02
19	41	100.0	24	31	100.0	19	ATM_HUMAN	1.03e-02
20	41	100.0	24	31	100.0	20	ATM_HUMAN	1.03e-02









[illegible][illegible]



[illegible]

RA SODLI J.D., SHIRLEY P., LILLY L.J., GILDER A., KELLEY J.M.,  
 RA WEISMAN S.P., FREEDMAN S.A., FRANKS J., HELLMER E.W., ELIASH H.L.,  
 RA CHERMAN V.E., HANNA M., MONTYEN E.J., PATTERSON M., REARDON R.D.,  
 RA FINE L.D., FRECHMAN J.L., FUHRMANN J.L., GEORGESEN R.S.M.,  
 RA GRIMM G.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,  
 RA VENTER J.C.,  
 RA "RA Nucleotide sequence and assembly of West-African  
 RA Influenza A virus", 1995,  
 RA Science 269:450-454 (1995),  
 RA 1. SIMILARITY: BELONGS TO THE SUI1 FAMILY.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. See [http://www.isb-sib.ch/announcements/](http://www.isb-sib.ch/announcements)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)  
 CC  
 DR EMBL: U32802; G1574156; -  
 DR TIGR: H1225; -  
 DR PROSITE: PS01118; SUI1; 1.  
 DR PFAM: PF01253; SUI1; 1.  
 KW HYPOTHETICAL PROTEIN: PROTEIN BIOSYNTHESIS, TRANSLATION REGULATION  
 SQ SEQUENCE 106 AA; 11226 MW; 8551932 CRC32;  
 Query Match 106.04 Score 31.0 E-1 Length 106.  
 Best Local Similarity 100.00; Pred. No. 1.03e-02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 22 VWRP 24  
 1111  
 QY 1 VWRP 4  
 RESULT 8  
 ID YFDO\_ECOLI STANDARD; FR1: 122 AA  
 AC P75311; P75343;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DI 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DI 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 14.4 KD PROTEIN IN INTX-DSDC INTERGENIC REGION.  
 GN YFDO.  
 OS ESCHERICHIA COLI.  
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
 OC ESCHERICHIA.  
 RN [1]  
 PP SEQUENCE FROM N.A.  
 RX STRAIN=K12; MG1655;  
 FX MEDLINE: 9125517;  
 RA BLATTNER F.R., PLUNKETT G., III, BLOCH C.A., PERNA N.T., BURLAND V.,  
 RA PIFFET M., CHILADO-VIRIS J., GLASNER F.D., POTTE J.K., MAYHEW G.F.,  
 RA GREIG J., LAVS N.W., KERRAULTER H.A., JELLEN M.A., ROSE J.J.,  
 RA MAUR R., SHAO Y.,  
 RA "The complete genome sequence of Escherichia coli F-10",  
 RA SCIENT. 277 1433-1474 (1997),  
 CC (2)  
 PP SEQUENCE FROM N.A.  
 RX STRAIN=K12;  
 RA AIBA H., BABA T., FUJITA K., HAYASHI K., HONTO A., HORIOCHI T.,  
 RA IEMOTO F., INATA T., ISINO K., KIMURA S., ICHIHARA T., KANAY K.,  
 RA KASAI H., KASHIMOTO K., KIM S., KIMURA S., KITAGAWA M.,  
 RA KITAKAWA M., MAKINO K., MASUDA S., MIKI T., MIYOSHI K., MORI H.,  
 RA MOTOHARA K., NAKAMURA Y., NASHIMOTO H., NISHIO Y., OSHIMA Y.,  
 RA SAITO N., SAMPEI G., SEKI Y., TAGAMI H., IASEMOTO K., WADA C.,  
 RA YAMAMOTO Y., YANG M.,  
 RA YAMAMOTO Y.,  
 RA SUBMITTED (JAN-1997) TO EMBL/GENBANK/CCAT DATA BANKS  
 CC 1. SIMILARITY: BELONGS TO THE PHAGE LAMBA REPLICATION PROTEIN 5 AND 10  
 CC BACTERIOPHAGE PH140 REPLICATION PROTEIN 1.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. See <http://www.isb-sib.ch/announcements/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)  
 CC  
 DR EMBL: AF000104; G1789700; -  
 DR FMR1; FMR1; 3174723; AUT\_INIT.  
 DR ECGENE1; EGI4137; YFDO.  
 RA HYDROPHETICAL PROTEIN:  
 SQ SEQUENCE 122 AA; 14431 MW; 18708684 CRC32;  
 Query Match 100.00 Score 31.0 E-1 Length 122.  
 Best Local Similarity 100.00; Pred. No. 1.03e-02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 22 VWRP 24  
 1111  
 QY 1 VWRP 4  
 RESULT 9  
 ID TRJ6\_ECOLI STANDARD; FR1: 122 AA  
 AC P17909;  
 DT 01-NOV-1990 (REL. 16, CREATED)  
 DI 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)  
 DI 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)  
 DE TRAJ PROTEIN (RELAX-SOME PROTEIN).  
 GN TRAJ.  
 OS ESCHERICHIA COLI.  
 OC PLASMID INCP-BETA RP4.  
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
 OC ESCHERICHIA.  
 RN [1]  
 PP SEQUENCE FROM N.A. AND SEQUENCE OF 1-5 AND 118-122.  
 RX MEDLINE: R030605;  
 RA ZIEGLIN G., FUERST J.P., LANKA E.,  
 RA "Traj protein of plasmid RP4 binds to a 19-base pair invert sequence  
 RA repetition within the transfer origin",  
 RA J. BIOL. CHEM. 264:11989-11994 (1989),  
 CC (2)  
 PP SEQUENCE FROM N.A.  
 RX STRAIN=HB101;  
 FX MEDLINE: 92190548;  
 RA ZIEGLIN G., FANSEGRAD W., SIRACE E., PALZER D., KREGER M., KRUET V.,  
 RA LANKA E.,  
 RA "Nucleotide sequence and organization of genes flanking the transfer  
 RA origin of promiscuous plasmid RP4",  
 RA J. BIOL. CHEM. 264:11995-12001 (1989),  
 CC (3)  
 PP SEQUENCE OF 1-7 FROM N.A.  
 RX MEDLINE: R0184510;  
 RA LANKA E., FURST J.P.,  
 RA "Nucleotide sequence of the plasmid transfer origin",  
 RA J. BIOL. CHEM. 264:11995-12001 (1989),  
 CC (4)  
 PP FUNCTION: TRANSFER OF PLASMID RP4 DURING BACTERIAL CONJUGATION  
 CC REQUIRES THE PLASMID-ENCODED TRAJ PROTEIN, WHICH BINDS TO A 19-  
 CC BASE PAIR INVERT SEQUENCE REPEITION WITHIN THE TRANSFER ORIGIN.  
 CC TRAJ PROTEIN IS BOUND TO ONLY ONE SIDE OF THE DNA HELIX. THIS  
 CC NUCLEOPROTEIN STRUCTURE IS THE INITIAL COMPLEX IN THE PATHWAY TO  
 CC ASSEMBLE A FUNCTIONAL RELAXOSOME.  
 CC 1. SURUNIT: MONOMER.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. See <http://www.isb-sib.ch/announcements/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)  
 CC  
 DR EMBL: X54452; G459364; -  
 DR EMBL: X54459; G45785; -



CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.embnet.fr/embnet/>  
 CC or send an email to [license@ebi-sib.ch](mailto:license@ebi-sib.ch).)

CC -----  
 CC EMBL: U12775: G540073: -  
 CC FMR1: U12770: G540074: JOINED  
 CC FMR1: U12774: G540075: JOINED  
 CC EMBL: L37019: G540046: -  
 CC DR MIM: 600201: -  
 CC SIGNAL: GLYCOPROTEIN: POLYMORPHISM.  
 CC FT SIGNAL: 1 32  
 CC FT CHAIN: 23 132  
 CC FT DOMAIN: 57 86  
 CC FT DOMAIN: 93 132  
 CC FT CARBOHYD: 22 29  
 CC FT VARIANT: 61 61  
 CC SQ SEQUENCE: 132 AA: 14515 WK: 5505862 CP032:

Query Match: 100.0% Score 31. DB 1. Length 132  
 Best Local Similarity 100.0% Pred. No. 103e-02  
 Matches 4. Conservative 0. Mismatches 0. Gaps 0.

DB 82 VWRP 95  
 QY 1 VWRP 4

RESULT 17  
 ID PK16-CHLUV STANDARD: PRI: 134 AA.  
 AC P55364:  
 DT 15-JUL-1994 (REL 35, CREATED)  
 DT 15-JUL-1998 (REL 36, LAST SEQUENCE UPDATE)  
 DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L16.  
 GN RPL16.  
 OS CHLORELLA VULGARIS.  
 OC CHLOROPLAST.  
 OC EUKARYOTA: VIRIDIPLANIA: CHLOROPHYTA: CHLOROPHYTA: CHLORELLALES:  
 CC CHLORELLACEAE: CHLORELLA.  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN: TAM 1111.  
 CC MEDLINE: 97303241.  
 CC RA WAKASUGI T, NAGAI T, KAPOR M, SUGITA M, ITO M, ITO S,  
 CC TSUBUKI T, NAKASHIMA K, TSUBUKI T, SUZUKI Y, HAMADA A, OHTA T,  
 CC INAMURA A, YOSHINAGA K, SUGIURA M.  
 CC "Complete nucleotide sequence of the chloroplast genome from the  
 CC green alga Chlorella vulgaris: the existence of genes possibly  
 CC involved in chloroplast division."  
 CC PROC Natl Acad Sci U S A 94:5567-5572(1997).  
 CC \* SIMILARITY: BELONGS TO THE L16P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.embnet.fr/embnet/>  
 CC or send an email to [license@ebi-sib.ch](mailto:license@ebi-sib.ch).)

CC -----  
 CC FMR1: AR001684: D1021583: -  
 CC PROSITE: PS00586: RIBOSOMAL\_L16\_1: 1  
 CC PROSITE: PS00701: RIBOSOMAL\_L16\_2: 1  
 CC KW RIBOSOMAL PROTEIN CHLOROPLAST  
 CC SEQUENCE 134 AA: 15136 WK: 352432CA CP032:

Query Match: 100.0% Score 31. DB 1. Length 134.  
 Best Local Similarity 100.0% Pred. No. 103e-02  
 Matches 4. Conservative 0. Mismatches 0. Gaps 0.

DB 95 VWRP 93

QY 1 VWRP 4  
 RESULT 14  
 ID HV15 MOUSE STANDARD: PRI: 136 AA.  
 AC P01759:  
 DT 21-JUL-1986 (REL 01, CREATED)  
 DT 21-JUL-1986 (REL 01, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1991 (REL 17, LAST ANNOTATION UPDATE)  
 DE IG HEAVY CHAIN PREDECESSOR V REGION (BOLL).  
 OS MUS MUSCULUS (MOUSE).  
 CC EUPARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA:  
 CC RODENTIA: SCIROGNATHI: MURIDAE: MURINAE, MUS.  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC MEDLINE: 92020242.  
 CC RA KNATHE M, P, LILJ P, NEWELL N, WARD E H, TUCKER P W, STEINER S,  
 CC KNATHE P P.  
 CC "Simultaneous expression of immunoglobulin mu and delta heavy chains  
 CC by a cloned B-cell hybridoma: a step in the VH gene is shared  
 CC by two adjacent CH genes."  
 CC EMBL: NAITL ACAL 531.0 S.A. 73.2222-3000(1982).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.embnet.fr/embnet/>  
 CC or send an email to [license@ebi-sib.ch](mailto:license@ebi-sib.ch).)

CC -----  
 CC FMR1: J06434: G195011: -  
 CC FMR1: A20421: HVMSB1:  
 CC FMR1: P00047: 19: 1.  
 CC IMMUNOGLOBULIN V REGION: SIGNAL.  
 CC SIGNAL 1 19  
 CC FT CHAIN 20 126  
 CC FT NONTER 136 136  
 CC SEQUENCE 136 AA: 15078 MW: 10491077 CP032:  
 CC Query Match: 100.0% Score 31. DB 1. Length 136:  
 CC Best Local Similarity 100.0% Pred. No. 103e-02:  
 CC Matches 4. Conservative 0. Mismatches 0. Gaps 0:

DB 30 VWRP 33  
 QY 1 VWRP 4

RESULT 15  
 ID SCHE STEHA STANDARD: PRI: 138 AA.  
 AC Q05352:  
 DT 01-FEB-1995 (REL 31, CREATED)  
 DT 01-FEB-1995 (REL 31, LAST SEQUENCE UPDATE)  
 DE SCHE STEHA (REL 32, LAST ANNOTATION UPDATE)  
 GN SCHE STEHA.  
 OS STEPTOMYCES HALSTEDII.  
 CC BACTERIA: FIRMICUTES: ACIDIMBACTERIA: ACIDIMBACTERIACE:  
 CC ACTINOMYCETALES: STEPTOMYCETALES: STEPTOMYCETACEAE: STEPTOMYCETES.  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN: NPT 2381:  
 CC MEDLINE: 93345807.  
 CC RA BLANCO G, BLAN P, PERLA A, MENDEL C, SALAS J A, CHAIER K P,  
 CC "Phylogenetic and DNA sequence analyses suggest an early  
 CC evolutionary divergence of related filamentous gene sets encoding  
 CC polyketide synthetases and secreted fragments in Streptomyces spp."  
 CC GENE 140:67-114(1993).  
 CC \*\*\* FAMILIARITY: BELONGS TO THE STEPTOMYCETACEAE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration



00 between the Swiss Institute of Bioinformatics and the EMBL outstation.  
 01 the European Bioinformatics Institute. There are no restrictions on its  
 02 use for non-profit institutions as long as its content is in no way  
 03 commercial and this statement is not removed, copied by and for commercial  
 04 purposes or put in a public domain (see: <http://www.ebi.ac.uk/press/pr000101.html>  
 05 or send an email to [listserv@ebi.ac.uk](mailto:listserv@ebi.ac.uk)).

06 -----  
 07 EMBL accession: D15432.1  
 08 EMBL accession: D15432.1  
 09 SEQUENCE: 146 AA; 15002 MW; 22888700 CRC32;

10 Query Match: 100.0%; Score 31; DR 1; Length 146;  
 11 Best local Similarity 100.0%; Pred. No. 1,036-02;  
 12 Matches 4; Conserved: 0; Mismatches 0; Indels 0; Gaps 0;

13 14 VWRP 47  
 14 15 VWRP 4

15 Sent to: [listserv@ebi.ac.uk](mailto:listserv@ebi.ac.uk) Fri Feb 25 13:07:00 2000  
 16 Subject: 47 seqs.











132	29	93.5	605	6.95e-02	ORF46A.	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
133	29	93.5	606	6.95e-02	HYPOETHICAL.	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
134	29	93.5	607	6.95e-02	ENVELOPE GLYCOPROTEIN.	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
135	29	93.5	608	6.95e-02	GENOME PARTIAL SEQUE	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
136	29	93.5	609	6.95e-02	PK31 BETA (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
137	29	93.5	610	6.95e-02	HYPOETHICAL 9.1 KD FR	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
138	29	93.5	611	6.95e-02	TAT ORF.	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
139	29	93.5	612	6.95e-02	CONSERVED PROTEIN.	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
140	29	93.5	613	6.95e-02	ENVELOPE GLYCOPROTEIN	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
141	29	93.5	614	6.95e-02	ENVELOPE GLYCOPROTEIN	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
142	29	93.5	615	6.95e-02	ENVELOPE GLYCOPROTEIN.	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
143	29	93.5	616	6.95e-02	VIRAL PROTEIN 1 (FRAGM	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
144	29	93.5	617	6.95e-02	CP31 PROTEIN.	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
145	29	93.5	618	6.95e-02	CHLOROPLAST PROTEINS (P	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
146	29	93.5	619	6.95e-02	ENVELOPE GLYCOPROTEIN	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
147	29	93.5	620	6.95e-02	CELL WALL PEPTIDE RICH	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
148	29	93.5	621	6.95e-02	VIRAL SHEDDING, V3	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
149	29	93.5	622	6.95e-02	MEAN PROTEIN.	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
150	29	93.5	623	6.95e-02	GLNB-LIKE PROTEIN.	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
151	29	93.5	624	6.95e-02	HYPOETHICAL PROTEIN (	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
152	29	93.5	625	6.95e-02	IMMUNOPROTEIN.	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
153	29	93.5	626	6.95e-02	ENVELOPE GLYCOPROTEIN	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
154	29	93.5	627	6.95e-02	ENVELOPE GLYCOPROTEIN	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
155	29	93.5	628	6.95e-02	ENVELOPE GLYCOPROTEIN	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
156	29	93.5	629	6.95e-02	ENVELOPE GLYCOPROTEIN	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
157	29	93.5	630	6.95e-02	TUNGSTON FORMYL METHAN	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
158	29	93.5	631	6.95e-02	TUNGSTON FORMYL METHAN	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
159	29	93.5	632	6.95e-02	TUNGSTON FORMYL METHAN	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
160	29	93.5	633	6.95e-02	ORF1	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
161	29	93.5	634	6.95e-02	SSD RIBOSOMAL PROTEIN	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
162	29	93.5	635	6.95e-02	ORF13.	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
163	29	93.5	636	6.95e-02	ORF11 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
164	29	93.5	637	6.95e-02	REIZING RELATED PROTEI	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
165	29	93.5	638	6.95e-02	15.5 KDA OLIGOMER.	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
166	29	93.5	639	6.95e-02	AAT (AAT) AND TEXE	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
167	29	93.5	640	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
168	29	93.5	641	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
169	29	93.5	642	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
170	29	93.5	643	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
171	29	93.5	644	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
172	29	93.5	645	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
173	29	93.5	646	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
174	29	93.5	647	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
175	29	93.5	648	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
176	29	93.5	649	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
177	29	93.5	650	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
178	29	93.5	651	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
179	29	93.5	652	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
180	29	93.5	653	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
181	29	93.5	654	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
182	29	93.5	655	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
183	29	93.5	656	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
184	29	93.5	657	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
185	29	93.5	658	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
186	29	93.5	659	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
187	29	93.5	660	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
188	29	93.5	661	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
189	29	93.5	662	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
190	29	93.5	663	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
191	29	93.5	664	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
192	29	93.5	665	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
193	29	93.5	666	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
194	29	93.5	667	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
195	29	93.5	668	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
196	29	93.5	669	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
197	29	93.5	670	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
198	29	93.5	671	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
199	29	93.5	672	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
200	29	93.5	673	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
201	29	93.5	674	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
202	29	93.5	675	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
203	29	93.5	676	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02
204	29	93.5	677	6.95e-02	ORF12 (FRAGMENT).	6.95e-02	6.95e-02	6.95e-02	VIF PROTEIN.	6.95e-02









11 11 VWRP 4  
 12 11 11  
 13 11 VWRP 4  
 14 11 11  
 15 11 VWRP 4  
 16 11 11  
 17 11 VWRP 4  
 18 11 11  
 19 11 VWRP 4  
 20 11 11  
 21 11 VWRP 4  
 22 11 11  
 23 11 VWRP 4  
 24 11 11  
 25 11 VWRP 4  
 26 11 11  
 27 11 VWRP 4  
 28 11 11  
 29 11 VWRP 4  
 30 11 11  
 31 11 VWRP 4  
 32 11 11  
 33 11 VWRP 4  
 34 11 11  
 35 11 VWRP 4  
 36 11 11  
 37 11 VWRP 4  
 38 11 11  
 39 11 VWRP 4  
 40 11 11  
 41 11 VWRP 4  
 42 11 11  
 43 11 VWRP 4  
 44 11 11  
 45 11 VWRP 4  
 46 11 11  
 47 11 VWRP 4  
 48 11 11  
 49 11 VWRP 4  
 50 11 11  
 51 11 VWRP 4  
 52 11 11  
 53 11 VWRP 4  
 54 11 11  
 55 11 VWRP 4  
 56 11 11  
 57 11 VWRP 4  
 58 11 11  
 59 11 VWRP 4  
 60 11 11  
 61 11 VWRP 4  
 62 11 11  
 63 11 VWRP 4  
 64 11 11  
 65 11 VWRP 4  
 66 11 11  
 67 11 VWRP 4  
 68 11 11  
 69 11 VWRP 4  
 70 11 11  
 71 11 VWRP 4  
 72 11 11  
 73 11 VWRP 4  
 74 11 11  
 75 11 VWRP 4  
 76 11 11  
 77 11 VWRP 4  
 78 11 11  
 79 11 VWRP 4  
 80 11 11  
 81 11 VWRP 4  
 82 11 11  
 83 11 VWRP 4  
 84 11 11  
 85 11 VWRP 4  
 86 11 11  
 87 11 VWRP 4  
 88 11 11  
 89 11 VWRP 4  
 90 11 11  
 91 11 VWRP 4  
 92 11 11  
 93 11 VWRP 4  
 94 11 11  
 95 11 VWRP 4  
 96 11 11  
 97 11 VWRP 4  
 98 11 11  
 99 11 VWRP 4  
 100 11 11

SO SEQUENCE 69 AA: 7941 MW: 14854550 CPG42  
 Query Match 100.0% Score 31: 10 23 1000's 100  
 Best Local Similarity 100.0% Prod. No. 2.050042  
 Matches 4: Conservative 0: Mismatches 0: Gaps 0  
 DB 35 VWRP 38  
 11 11  
 12 11 VWRP 4  
 13 11 11  
 14 11 VWRP 4  
 15 11 11  
 16 11 VWRP 4  
 17 11 11  
 18 11 VWRP 4  
 19 11 11  
 20 11 VWRP 4  
 21 11 11  
 22 11 VWRP 4  
 23 11 11  
 24 11 VWRP 4  
 25 11 11  
 26 11 VWRP 4  
 27 11 11  
 28 11 VWRP 4  
 29 11 11  
 30 11 VWRP 4  
 31 11 11  
 32 11 VWRP 4  
 33 11 11  
 34 11 VWRP 4  
 35 11 11  
 36 11 VWRP 4  
 37 11 11  
 38 11 VWRP 4  
 39 11 11  
 40 11 VWRP 4  
 41 11 11  
 42 11 VWRP 4  
 43 11 11  
 44 11 VWRP 4  
 45 11 11  
 46 11 VWRP 4  
 47 11 11  
 48 11 VWRP 4  
 49 11 11  
 50 11 VWRP 4  
 51 11 11  
 52 11 VWRP 4  
 53 11 11  
 54 11 VWRP 4  
 55 11 11  
 56 11 VWRP 4  
 57 11 11  
 58 11 VWRP 4  
 59 11 11  
 60 11 VWRP 4  
 61 11 11  
 62 11 VWRP 4  
 63 11 11  
 64 11 VWRP 4  
 65 11 11  
 66 11 VWRP 4  
 67 11 11  
 68 11 VWRP 4  
 69 11 11  
 70 11 VWRP 4  
 71 11 11  
 72 11 VWRP 4  
 73 11 11  
 74 11 VWRP 4  
 75 11 11  
 76 11 VWRP 4  
 77 11 11  
 78 11 VWRP 4  
 79 11 11  
 80 11 VWRP 4  
 81 11 11  
 82 11 VWRP 4  
 83 11 11  
 84 11 VWRP 4  
 85 11 11  
 86 11 VWRP 4  
 87 11 11  
 88 11 VWRP 4  
 89 11 11  
 90 11 VWRP 4  
 91 11 11  
 92 11 VWRP 4  
 93 11 11  
 94 11 VWRP 4  
 95 11 11  
 96 11 VWRP 4  
 97 11 11  
 98 11 VWRP 4  
 99 11 11  
 100 11 VWRP 4

```

Best Local Similarity 100.0%, Score 31, DB 14, Length 81;
Matches 4: Conservative 0, Mismatches 0, Indels 0, Gaps 0;

DB 6 VWRP 9
    ||||
QY 1 VWRP 4

RESULT 9
ID Q36468 PRELIMINARY: PRI: 78 AA
AC Q05459
DT 01-JAN-1998 (PREMBLREL 04, CREATED)
DI 01-JUL-1997 (PREMBLREL 04, LAST SEQUENCE UPDATE)
DE 01-JUL-1997 (PREMBLREL 04, LAST ANNOTATION UPDATE)
OS RHIZOBIUM TROPICI
OC RHIZOBIACEAE, RHIZOBIUM
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CIAT999;
FA CHATTERJEE R., TSAREV S., PILOT J., COURSAGET P., EMERSON S.,
  PURCELL R.;
EMBL: AF010425; G1923317; -.
FT NON-TER 78 78
SQ SEQUENCE 78 AA: 8417 MW: 887167.00
Query Match 100.0%, Score 31, DB 14, Length 78;
Best Local Similarity 100.0%, Pref No 2,050-02;
Matches 4: Conservative 0, Mismatches 0, Indels 0, Gaps 0;

DB 44 VWRP 47
    ||||
QY 1 VWRP 4

RESULT 10
ID Q36602 PRELIMINARY: PRI: 81 AA
AC Q36604
DT 01-JAN-1998 (PREMBLREL 05, CREATED)
DI 01-JAN-1998 (PREMBLREL 05, LAST SEQUENCE UPDATE)
DE 01-JAN-1998 (PREMBLREL 05, LAST ANNOTATION UPDATE)
OS STRAIN KIRGIZIA (OSH)
OC VIRUSAE, SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE, CALICIVIRIDAE
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KIRGIZIA (OSH);
FA CHATTERJEE R., TSAREV S., PILOT J., COURSAGET P., EMERSON S.,
  PURCELL R.;
EMBL: AF010425; G1923317; -.
FT NON-TER 81 81
SQ SEQUENCE 81 AA: 9588 MW: 907078.00
Query Match 100.0%, Score 31, DB 14, Length 81;
Best Local Similarity 100.0%, Pref No 2,050-02;
Matches 4: Conservative 0, Mismatches 0, Indels 0, Gaps 0;

DB 1 VWRP 2
    ||||
QY 1 VWRP 4

RESULT 11
ID Q36603 PRELIMINARY: PRI: 81 AA
AC Q36604
DT 01-JAN-1998 (PREMBLREL 05, CREATED)
DI 01-JAN-1998 (PREMBLREL 05, LAST SEQUENCE UPDATE)
DE 01-JAN-1998 (PREMBLREL 05, LAST ANNOTATION UPDATE)
OS STRAIN KIRGIZIA (OSH)
OC VIRUSAE, SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE, CALICIVIRIDAE
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KIRGIZIA (OSH);
FA CHATTERJEE R., TSAREV S., PILOT J., COURSAGET P., EMERSON S.,
  PURCELL R.;
EMBL: AF010425; G1923317; -.
FT NON-TER 81 81
SQ SEQUENCE 81 AA: 9588 MW: 907078.00
Query Match 100.0%, Score 31, DB 14, Length 81;
Best Local Similarity 100.0%, Pref No 2,050-02;
Matches 4: Conservative 0, Mismatches 0, Indels 0, Gaps 0;

DB 1 VWRP 2
    ||||
QY 1 VWRP 4

RESULT 12
ID Q36603 PRELIMINARY: PRI: 81 AA
AC Q36604
DT 01-JAN-1998 (PREMBLREL 05, CREATED)
DI 01-JAN-1998 (PREMBLREL 05, LAST SEQUENCE UPDATE)
DE 01-JAN-1998 (PREMBLREL 05, LAST ANNOTATION UPDATE)
OS STRAIN KIRGIZIA (OSH)
OC VIRUSAE, SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE, CALICIVIRIDAE
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KIRGIZIA (OSH);
FA CHATTERJEE R., TSAREV S., PILOT J., COURSAGET P., EMERSON S.,
  PURCELL R.;
EMBL: AF010425; G1923317; -.
FT NON-TER 81 81
SQ SEQUENCE 81 AA: 9588 MW: 907078.00
Query Match 100.0%, Score 31, DB 14, Length 81;
Best Local Similarity 100.0%, Pref No 2,050-02;
Matches 4: Conservative 0, Mismatches 0, Indels 0, Gaps 0;

DB 2 VWRP 5
    ||||
QY 1 VWRP 4

RESULT 13
ID Q36601 PRELIMINARY: PRI: 81 AA
AC Q36601
DT 01-JAN-1998 (PREMBLREL 05, CREATED)
DI 01-JAN-1998 (PREMBLREL 05, LAST SEQUENCE UPDATE)
DE 01-JAN-1998 (PREMBLREL 05, LAST ANNOTATION UPDATE)
OS STRAIN KIRGIZIA (OSH)
OC VIRUSAE, SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE, CALICIVIRIDAE
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KIRGIZIA (OSH);
FA CHATTERJEE R., TSAREV S., PILOT J., COURSAGET P., EMERSON S.,
  PURCELL R.;
EMBL: AF010425; G1923317; -.
FT NON-TER 81 81
SQ SEQUENCE 81 AA: 9588 MW: 907078.00
Query Match 100.0%, Score 31, DB 14, Length 81;
Best Local Similarity 100.0%, Pref No 2,050-02;
Matches 4: Conservative 0, Mismatches 0, Indels 0, Gaps 0;

DB 2 VWRP 5
    ||||
QY 1 VWRP 4

```

```

STRAIN KIRGIZIA (TASHKENT) ORF-1 (FRAGMENT);
HEPATITIS E VIRUS (HEV);
VIRUSES, SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE, CALICIVIRIDAE
OC CALICIVIRUS
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-UZBEKISTAN (TASHKENT);
FA CHATTERJEE R., TSAREV S., PILOT J., COURSAGET P., EMERSON S.,
  PURCELL R.;
EMBL: AF010425; G1923317; -.
FT NON-TER 81 81
SQ SEQUENCE 81 AA: 9588 MW: 907078.00
Query Match 100.0%, Score 31, DB 14, Length 81;
Best Local Similarity 100.0%, Pref No 2,050-02;
Matches 4: Conservative 0, Mismatches 0, Indels 0, Gaps 0;

DB 2 VWRP 5
    ||||
QY 1 VWRP 4

RESULT 12
ID Q36603 PRELIMINARY: PRI: 81 AA
AC Q36604
DT 01-JAN-1998 (PREMBLREL 05, CREATED)
DI 01-JAN-1998 (PREMBLREL 05, LAST SEQUENCE UPDATE)
DE 01-JAN-1998 (PREMBLREL 05, LAST ANNOTATION UPDATE)
OS STRAIN KIRGIZIA (TASHKENT) ORF-1 (FRAGMENT);
HEPATITIS E VIRUS (HEV);
VIRUSES, SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE, CALICIVIRIDAE
OC CALICIVIRUS
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MOROCCO/F12;
FA CHATTERJEE R., TSAREV S., PILOT J., COURSAGET P., EMERSON S.,
  PURCELL R.;
EMBL: AF010425; G1923317; -.
FT NON-TER 81 81
SQ SEQUENCE 81 AA: 9588 MW: 907078.00
Query Match 100.0%, Score 31, DB 14, Length 81;
Best Local Similarity 100.0%, Pref No 2,050-02;
Matches 4: Conservative 0, Mismatches 0, Indels 0, Gaps 0;

DB 2 VWRP 5
    ||||
QY 1 VWRP 4

RESULT 13
ID Q36601 PRELIMINARY: PRI: 81 AA
AC Q36601
DT 01-JAN-1998 (PREMBLREL 05, CREATED)
DI 01-JAN-1998 (PREMBLREL 05, LAST SEQUENCE UPDATE)
DE 01-JAN-1998 (PREMBLREL 05, LAST ANNOTATION UPDATE)
OS STRAIN KIRGIZIA (TASHKENT) ORF-1 (FRAGMENT);
HEPATITIS F VIRUS (HFV)
VIRUSES, SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE, CALICIVIRIDAE
OC CALICIVIRUS
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MOROCCO/F23;
FA CHATTERJEE R., TSAREV S., PILOT J., COURSAGET P., EMERSON S.,
  PURCELL R.;
EMBL: AF010425; G1923317; -.
FT NON-TER 81 81
SQ SEQUENCE 81 AA: 9588 MW: 907078.00
Query Match 100.0%, Score 31, DB 14, Length 81;
Best Local Similarity 100.0%, Pref No 2,050-02;
Matches 4: Conservative 0, Mismatches 0, Indels 0, Gaps 0;

DB 2 VWRP 5
    ||||
QY 1 VWRP 4

```





# WARNING

Release 3.0A John F. Collins, Bioinformatics Research Unit,

University of Edinburgh, University of Edinburgh, U.K.

Distribution rights by Oxford Molecular Ltd

Warning: If you find a protein database search using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

gives a result, please check the database search, using Smith-Waterman algorithm

17 37 94.9 1157 5 R25957 2.98e-02  
18 37 94.9 1227 29 W54921 2.98e-02

Note: Postfixes of 1 and 2 indicate the 1st and 2nd similarity from the 1st and 2nd similarity 1 and 2.

## ALIGNMENTS

RESULT 1  
12 W10102 standard: protein: 130 AA.  
AC W10102:  
DE Human agouti signalling protein.  
KW Agouti signalling protein: ASP, depigmenting activity; cosmeceutic  
KW hyperpigmentary condition; melanin pigmentation; skin; keratinocyte  
KW post-inflammatory hyperpigmentation; wound healing; melanogenesis  
KW vitiligo; leucoderma; albinism; hair greying.  
QS Homo sapiens.  
PH Key Location/Qualifiers  
ET peptide 1-21  
ET region 42 /note: "Signal sequence"  
ET region 86 /note: "Start of basic region."  
ET region 86 /note: "Start of cysteine-rich motif"

PR W09793892.A2.  
PD 09-JAN-1997.  
PF 21-JUN-1995: U10695.  
PR 23-JUN-1995: US-000436.  
PA (US) US DEPT HEALTH & HUMAN SERVICES.  
PI Hearing VI:  
DR WPI: 97-087323/09.  
PR New agouti signal protein peptide(s) and nucleic acids useful for  
PI altering melanin production for treating skin melanoma phenotype  
PI spots, skin; keratosis or vitiligo  
PS claim 6: Page 8-9: 67pp: English.  
CC The sequences given in W10101-29 are biologically active peptides  
CC and fragments of the agouti signalling protein (ASP) which have  
CC depigmenting activity. These peptides are useful for treatment of various  
CC and for clinical application in the prevention or treatment of various  
CC hyperpigmentary conditions and diseases such as melanoma pigmentation  
CC spots, skin; keratosis and post-inflammatory hyperpigmentation such as  
CC occurs at sites of wound healing. They can also be used to provide  
CC enhanced melanogenesis for treating e.g. vitiligo, leucoderma, some  
CC forms of albinism and hair greying.  
SQ Sequence 130 AA.

Query Match 100.0% Score 39: DR 24: Ident 14:  
Best Local Similarity 100.0% Pred. No: 1,680-02:  
Matches 5: Conservative 0: Mismatches 0: Indels 0: Gaps 0

DB 82 VRRPP 44  
CY 1 VRRPP 5

RESULT 2  
10 R10276 standard: protein: 171 AA.  
AC R10276:  
DE 04-APR-1991 (first entry)  
DE Simian immunodeficiency virus vif protein.  
DE SIV-2: HIV, AIDS; variable; RSH104.  
QS Simian immunodeficiency virus.  
PR 02295587-A.  
PD 11-DEC-1990.  
PF 11-MAY-1989: 12129.  
PR 11-MAY-1989: JF 11129.  
PA (TOFU) TONEN CORP.  
PR W10102: 11129/24.  
PR N-PSB: Q10203.  
PR New complementary DNA to RNA gene of simian immunodeficiency  
PR virus, used for preparing vaccine and diagnosis agent for AIDS  
PS claim 2: Fig 4: 14pp: Japanese.

Summary: Mean 33.463, Variance 34.470, scale 0.388

Note: W is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARY

Rank	No.	Score	Entry	Seq. #	Start	End	Description	Pred. No.
1	1	49	W10102	130	23	10102	Human agouti signalling	1,680-02
2	2	49	R10276	171	2	10276	Simian immunodeficiency	1,680-02
3	3	49	W10102	130	28	10102	PSK (WPK) protein 1	1,680-02
4	4	49	W10102	130	28	10102	Ruthenium Staphylococ	1,680-02
5	5	49	W10102	130	28	10102	Plasma factor 1	1,680-02
6	6	49	W10102	130	28	10102	Sequence translated f	1,680-02
7	7	49	W10102	130	28	10102	Plasma factor 1	1,680-02
8	8	49	W10102	130	28	10102	Polypeptide fragment	1,680-02
9	9	49	W10102	130	28	10102	Polypeptide fragment	1,680-02
10	10	49	W10102	130	28	10102	Polypeptide fragment	1,680-02
11	11	49	W10102	130	28	10102	Polypeptide fragment	1,680-02
12	12	49	W10102	130	28	10102	Polypeptide fragment	1,680-02
13	13	49	W10102	130	28	10102	Polypeptide fragment	1,680-02
14	14	49	W10102	130	28	10102	Polypeptide fragment	1,680-02
15	15	49	W10102	130	28	10102	Polypeptide fragment	1,680-02
16	16	49	W10102	130	28	10102	Polypeptide fragment	1,680-02
17	17	49	W10102	130	28	10102	Polypeptide fragment	1,680-02
18	18	49	W10102	130	28	10102	Polypeptide fragment	1,680-02
19	19	49	W10102	130	28	10102	Polypeptide fragment	1,680-02
20	20	49	W10102	130	28	10102	Polypeptide fragment	1,680-02

CC SIV proteins may be produced from an Fv-100 expression system  
 CC transformed with plasmid pSMH103.  
 CC These may be used in diagnosis, treatment, and development of a vaccine  
 CC against AIDS, as the sequence has the same structure as HIV-1 (but  
 CC lacking the vpx gene).  
 CC Sequence 172 AA.

Query Match 100.0% Score 39. DB 21. Length 172.  
 Best Local Similarity 100.0% Pred No. 1.69e-02.  
 Matches 5. Conservation 0. Mismatches 0. Indels 0. Gaps 0.

DE 25 WRPP 99  
 |||||  
 QY 1 WRPP 5

RESULT 3  
 ID W37363 standard; Protein; 320 AA.  
 AC W37363  
 DT 27-MAR-1998 (first entry)  
 DE PSB (W37363) protein involved in bacterial synthesis and assembly  
 FW O antigen; psb; W37363; lipopolysaccharide; infection; diagnosis;  
 KW antibody; dehydratase.  
 OS Pseudomonas aeruginosa PA01.  
 FN W04741234.A2.  
 PD 06-NOV-1997.  
 RF 30-APR-1997; CA0295.  
 FR 27-FEB-1997; US-039473.  
 FR 30-APR-1996; US-015510.  
 PA (UYGU) UNIV GUELPH.  
 FI Barrows, J., Chatter, D., De Foevit, T., Lam, J.S.  
 DR WPI: 97-549236/50.  
 CR R-5528; 727221.  
 PT Pseudomonas aeruginosa B-band lipopolysaccharide gene cluster -  
 PI used for diagnosis of P. aeruginosa infection  
 PS Claim 8: Page 123; 195pp; English.  
 CC This sequence comprises PsbK (W37363), a Pseudomonas aeruginosa PA01  
 CC protein with dehydratase activity. W37363, PsbA (W37363), PsbB  
 CC (W37363), PsbC (W37363), PsbD (W37363), PsbE (W37363), PsbF  
 CC (W37363), PsbG (W37363), PsbH (W37363), PsbI (W37363), PsbJ  
 CC (W37363), PsbK (W37363), PsbL (W37363), PsbM (W37363), PsbN  
 CC (W37363), PsbO (W37363), PsbP (W37363), PsbQ (W37363), PsbR  
 CC (W37363), PsbS (W37363), PsbT (W37363), PsbU (W37363), PsbV  
 CC (W37363), PsbW (W37363), PsbX (W37363), PsbY (W37363), PsbZ  
 CC (W37363), PsbAA (W37363), PsbAB (W37363), PsbAC (W37363),  
 CC (W37363), PsbAD (W37363), PsbAE (W37363), PsbAF (W37363),  
 CC (W37363), PsbAG (W37363), PsbAH (W37363), PsbAI (W37363),  
 CC (W37363), PsbAJ (W37363), PsbAK (W37363), PsbAL (W37363),  
 CC (W37363), PsbAM (W37363), PsbAN (W37363), PsbAO (W37363),  
 CC (W37363), PsbAP (W37363), PsbAQ (W37363), PsbAR (W37363),  
 CC (W37363), PsbAS (W37363), PsbAT (W37363), PsbAU (W37363),  
 CC (W37363), PsbAV (W37363), PsbAW (W37363), PsbAX (W37363),  
 CC (W37363), PsbAY (W37363), PsbAZ (W37363), PsbBA (W37363),  
 CC (W37363), PsbBB (W37363), PsbBC (W37363), PsbBD (W37363),  
 CC (W37363), PsbBE (W37363), PsbBF (W37363), PsbBG (W37363),  
 CC (W37363), PsbBH (W37363), PsbBI (W37363), PsbBJ (W37363),  
 CC (W37363), PsbBK (W37363), PsbBL (W37363), PsbBM (W37363),  
 CC (W37363), PsbBN (W37363), PsbBO (W37363), PsbBP (W37363),  
 CC (W37363), PsbBQ (W37363), PsbBR (W37363), PsbBS (W37363),  
 CC (W37363), PsbBT (W37363), PsbBU (W37363), PsbBV (W37363),  
 CC (W37363), PsbBW (W37363), PsbBX (W37363), PsbBY (W37363),  
 CC (W37363), PsbBZ (W37363), PsbCA (W37363), PsbCB (W37363),  
 CC (W37363), PsbCC (W37363), PsbCD (W37363), PsbCE (W37363),  
 CC (W37363), PsbCF (W37363), PsbCG (W37363), PsbCH (W37363),  
 CC (W37363), PsbCI (W37363), PsbCJ (W37363), PsbCK (W37363),  
 CC (W37363), PsbCL (W37363), PsbCM (W37363), PsbCN (W37363),  
 CC (W37363), PsbCO (W37363), PsbCP (W37363), PsbCQ (W37363),  
 CC (W37363), PsbCR (W37363), PsbCS (W37363), PsbCT (W37363),  
 CC (W37363), PsbCU (W37363), PsbCV (W37363), PsbCW (W37363),  
 CC (W37363), PsbCX (W37363), PsbCY (W37363), PsbCZ (W37363),  
 CC (W37363), PsbDA (W37363), PsbDB (W37363), PsbDC (W37363),  
 CC (W37363), PsbDE (W37363), PsbDF (W37363), PsbDG (W37363),  
 CC (W37363), PsbDH (W37363), PsbDI (W37363), PsbDJ (W37363),  
 CC (W37363), PsbDK (W37363), PsbDL (W37363), PsbDM (W37363),  
 CC (W37363), PsbDN (W37363), PsbDO (W37363), PsbDP (W37363),  
 CC (W37363), PsbDQ (W37363), PsbDR (W37363), PsbDS (W37363),  
 CC (W37363), PsbDT (W37363), PsbDU (W37363), PsbDV (W37363),  
 CC (W37363), PsbDW (W37363), PsbDX (W37363), PsbDY (W37363),  
 CC (W37363), PsbDZ (W37363), PsbEA (W37363), PsbEB (W37363),  
 CC (W37363), PsbEC (W37363), PsbED (W37363), PsbEE (W37363),  
 CC (W37363), PsbEF (W37363), PsbEG (W37363), PsbEH (W37363),  
 CC (W37363), PsbEI (W37363), PsbEJ (W37363), PsbEK (W37363),  
 CC (W37363), PsbEL (W37363), PsbEM (W37363), PsbEN (W37363),  
 CC (W37363), PsbEO (W37363), PsbEP (W37363), PsbEQ (W37363),  
 CC (W37363), PsbER (W37363), PsbES (W37363), PsbET (W37363),  
 CC (W37363), PsbEU (W37363), PsbEV (W37363), PsbEW (W37363),  
 CC (W37363), PsbEX (W37363), PsbEY (W37363), PsbEZ (W37363),  
 CC (W37363), PsbFA (W37363), PsbFB (W37363), PsbFC (W37363),  
 CC (W37363), PsbFE (W37363), PsbFF (W37363), PsbFG (W37363),  
 CC (W37363), PsbFH (W37363), PsbFI (W37363), PsbFJ (W37363),  
 CC (W37363), PsbFK (W37363), PsbFL (W37363), PsbFM (W37363),  
 CC (W37363), PsbFN (W37363), PsbFO (W37363), PsbFP (W37363),  
 CC (W37363), PsbFQ (W37363), PsbFR (W37363), PsbFS (W37363),  
 CC (W37363), PsbFT (W37363), PsbFU (W37363), PsbFV (W37363),  
 CC (W37363), PsbFW (W37363), PsbFX (W37363), PsbFY (W37363),  
 CC (W37363), PsbFZ (W37363), PsbGA (W37363), PsbGB (W37363),  
 CC (W37363), PsbGC (W37363), PsbGD (W37363), PsbGE (W37363),  
 CC (W37363), PsbGF (W37363), PsbGH (W37363), PsbGI (W37363),  
 CC (W37363), PsbGJ (W37363), PsbGK (W37363), PsbGL (W37363),  
 CC (W37363), PsbGM (W37363), PsbGN (W37363), PsbGO (W37363),  
 CC (W37363), PsbGP (W37363), PsbGQ (W37363), PsbGR (W37363),  
 CC (W37363), PsbGS (W37363), PsbGT (W37363), PsbGU (W37363),  
 CC (W37363), PsbGV (W37363), PsbGW (W37363), PsbGX (W37363),  
 CC (W37363), PsbGY (W37363), PsbGZ (W37363), PsbHA (W37363),  
 CC (W37363), PsbHB (W37363), PsbHC (W37363), PsbHD (W37363),  
 CC (W37363), PsbHE (W37363), PsbHF (W37363), PsbHG (W37363),  
 CC (W37363), PsbHH (W37363), PsbHI (W37363), PsbHJ (W37363),  
 CC (W37363), PsbHK (W37363), PsbHL (W37363), PsbHM (W37363),  
 CC (W37363), PsbHN (W37363), PsbHO (W37363), PsbHP (W37363),  
 CC (W37363), PsbHQ (W37363), PsbHR (W37363), PsbHS (W37363),  
 CC (W37363), PsbHT (W37363), PsbHU (W37363), PsbHV (W37363),  
 CC (W37363), PsbHW (W37363), PsbHX (W37363), PsbHY (W37363),  
 CC (W37363), PsbHZ (W37363), PsbIA (W37363), PsbIB (W37363),  
 CC (W37363), PsbIC (W37363), PsbID (W37363), PsbIE (W37363),  
 CC (W37363), PsbIF (W37363), PsbIG (W37363), PsbIH (W37363),  
 CC (W37363), PsbIJ (W37363), PsbIK (W37363), PsbIL (W37363),  
 CC (W37363), PsbIM (W37363), PsbIN (W37363), PsbIO (W37363),  
 CC (W37363), PsbIP (W37363), PsbIQ (W37363), PsbIR (W37363),  
 CC (W37363), PsbIS (W37363), PsbIT (W37363), PsbIU (W37363),  
 CC (W37363), PsbIV (W37363), PsbIW (W37363), PsbIX (W37363),  
 CC (W37363), PsbIY (W37363), PsbIZ (W37363), PsbJA (W37363),  
 CC (W37363), PsbJB (W37363), PsbJC (W37363), PsbJD (W37363),  
 CC (W37363), PsbJE (W37363), PsbJF (W37363), PsbJG (W37363),  
 CC (W37363), PsbJH (W37363), PsbJI (W37363), PsbJJ (W37363),  
 CC (W37363), PsbJK (W37363), PsbJL (W37363), PsbJM (W37363),  
 CC (W37363), PsbJN (W37363), PsbJO (W37363), PsbJP (W37363),  
 CC (W37363), PsbJQ (W37363), PsbJR (W37363), PsbJS (W37363),  
 CC (W37363), PsbJT (W37363), PsbJU (W37363), PsbJV (W37363),  
 CC (W37363), PsbJW (W37363), PsbJX (W37363), PsbJY (W37363),  
 CC (W37363), PsbJZ (W37363), PsbKA (W37363), PsbKB (W37363),  
 CC (W37363), PsbKC (W37363), PsbKD (W37363), PsbKE (W37363),  
 CC (W37363), PsbKF (W37363), PsbKG (W37363), PsbKH (W37363),  
 CC (W37363), PsbKI (W37363), PsbKJ (W37363), PsbKL (W37363),  
 CC (W37363), PsbKM (W37363), PsbKN (W37363), PsbKO (W37363),  
 CC (W37363), PsbKP (W37363), PsbKQ (W37363), PsbKR (W37363),  
 CC (W37363), PsbKS (W37363), PsbKT (W37363), PsbKU (W37363),  
 CC (W37363), PsbKV (W37363), PsbKW (W37363), PsbKX (W37363),  
 CC (W37363), PsbKY (W37363), PsbKZ (W37363), PsbLA (W37363),  
 CC (W37363), PsbLB (W37363), PsbLC (W37363), PsbLD (W37363),  
 CC (W37363), PsbLE (W37363), PsbLF (W37363), PsbLG (W37363),  
 CC (W37363), PsbLH (W37363), PsbLI (W37363), PsbLJ (W37363),  
 CC (W37363), PsbLK (W37363), PsbLL (W37363), PsbLM (W37363),  
 CC (W37363), PsbLN (W37363), PsbLO (W37363), PsbLP (W37363),  
 CC (W37363), PsbLQ (W37363), PsbLR (W37363), PsbLS (W37363),  
 CC (W37363), PsbLT (W37363), PsbLU (W37363), PsbLV (W37363),  
 CC (W37363), PsbLW (W37363), PsbLX (W37363), PsbLY (W37363),  
 CC (W37363), PsbLZ (W37363), PsbMA (W37363), PsbMB (W37363),  
 CC (W37363), PsbMC (W37363), PsbMD (W37363), PsbME (W37363),  
 CC (W37363), PsbMF (W37363), PsbMG (W37363), PsbMH (W37363),  
 CC (W37363), PsbMI (W37363), PsbMJ (W37363), PsbMK (W37363),  
 CC (W37363), PsbML (W37363), PsbMM (W37363), PsbMN (W37363),  
 CC (W37363), PsbMO (W37363), PsbMP (W37363), PsbMQ (W37363),  
 CC (W37363), PsbMR (W37363), PsbMS (W37363), PsbMT (W37363),  
 CC (W37363), PsbMU (W37363), PsbMV (W37363), PsbMW (W37363),  
 CC (W37363), PsbMX (W37363), PsbMY (W37363), PsbMZ (W37363),  
 CC (W37363), PsbNA (W37363), PsbNB (W37363), PsbNC (W37363),  
 CC (W37363), PsbND (W37363), PsbNE (W37363), PsbNF (W37363),  
 CC (W37363), PsbNG (W37363), PsbNH (W37363), PsbNJ (W37363),  
 CC (W37363), PsbNK (W37363), PsbNL (W37363), PsbNM (W37363),  
 CC (W37363), PsbNN (W37363), PsbNO (W37363), PsbNP (W37363),  
 CC (W37363), PsbNQ (W37363), PsbNR (W37363), PsbNS (W37363),  
 CC (W37363), PsbNT (W37363), PsbNU (W37363), PsbNV (W37363),  
 CC (W37363), PsbNW (W37363), PsbNX (W37363), PsbNY (W37363),  
 CC (W37363), PsbNZ (W37363), PsbOA (W37363), PsbOB (W37363),  
 CC (W37363), PsbOC (W37363), PsbOD (W37363), PsbOE (W37363),  
 CC (W37363), PsbOF (W37363), PsbOG (W37363), PsbOH (W37363),  
 CC (W37363), PsbOI (W37363), PsbOJ (W37363), PsbOK (W37363),  
 CC (W37363), PsbOL (W37363), PsbOM (W37363), PsbON (W37363),  
 CC (W37363), PsbOO (W37363), PsbOP (W37363), PsbOQ (W37363),  
 CC (W37363), PsbOR (W37363), PsbOS (W37363), PsbOT (W37363),  
 CC (W37363), PsbOU (W37363), PsbOV (W37363), PsbOW (W37363),  
 CC (W37363), PsbOX (W37363), PsbOY (W37363), PsbOZ (W37363),  
 CC (W37363), PsbPA (W37363), PsbPB (W37363), PsbPC (W37363),  
 CC (W37363), PsbPD (W37363), PsbPE (W37363), PsbPF (W37363),  
 CC (W37363), PsbPG (W37363), PsbPH (W37363), PsbPI (W37363),  
 CC (W37363), PsbPJ (W37363), PsbPK (W37363), PsbPL (W37363),  
 CC (W37363), PsbPM (W37363), PsbPN (W37363), PsbPO (W37363),  
 CC (W37363), PsbPP (W37363), PsbPQ (W37363), PsbPR (W37363),  
 CC (W37363), PsbPS (W37363), PsbPT (W37363), PsbPU (W37363),  
 CC (W37363), PsbPV (W37363), PsbPW (W37363), PsbPX (W37363),  
 CC (W37363), PsbPY (W37363), PsbPZ (W37363), PsbQA (W37363),  
 CC (W37363), PsbQB (W37363), PsbQC (W37363), PsbQD (W37363),  
 CC (W37363), PsbQE (W37363), PsbQF (W37363), PsbQG (W37363),  
 CC (W37363), PsbQH (W37363), PsbQI (W37363), PsbQJ (W37363),  
 CC (W37363), PsbQK (W37363), PsbQL (W37363), PsbQM (W37363),  
 CC (W37363), PsbQN (W37363), PsbQO (W37363), PsbQP (W37363),  
 CC (W37363), PsbQQ (W37363), PsbQR (W37363), PsbQS (W37363),  
 CC (W37363), PsbQT (W37363), PsbQU (W37363), PsbQV (W37363),  
 CC (W37363), PsbQW (W37363), PsbQX (W37363), PsbQY (W37363),  
 CC (W37363), PsbQZ (W37363), PsbRA (W37363), PsbRB (W37363),  
 CC (W37363), PsbRC (W37363), PsbRD (W37363), PsbRE (W37363),  
 CC (W37363), PsbRF (W37363), PsbRG (W37363), PsbRH (W37363),  
 CC (W37363), PsbRI (W37363), PsbRJ (W37363), PsbRK (W37363),  
 CC (W37363), PsbRL (W37363), PsbRM (W37363), PsbRN (W37363),  
 CC (W37363), PsbRO (W37363), PsbRP (W37363), PsbRQ (W37363),  
 CC (W37363), PsbRR (W37363), PsbRS (W37363), PsbRT (W37363),  
 CC (W37363), PsbRU (W37363), PsbRV (W37363), PsbRW (W37363),  
 CC (W37363), PsbRX (W37363), PsbRY (W37363), PsbRZ (W37363),  
 CC (W37363), PsbSA (W37363), PsbSB (W37363), PsbSC (W37363),  
 CC (W37363), PsbSD (W37363), PsbSE (W37363), PsbSF (W37363),  
 CC (W37363), PsbSG (W37363), PsbSH (W37363), PsbSI (W37363),  
 CC (W37363), PsbSJ (W37363), PsbSK (W37363), PsbSL (W37363),  
 CC (W37363), PsbSM (W37363), PsbSN (W37363), PsbSO (W37363),  
 CC (W37363), PsbSP (W37363), PsbSQ (W37363), PsbSR (W37363),  
 CC (W37363), PsbSS (W37363), PsbST (W37363), PsbSU (W37363),  
 CC (W37363), PsbSV (W37363), PsbSW (W37363), PsbSX (W37363),  
 CC (W37363), PsbSY (W37363), PsbSZ (W37363), PsbTA (W37363),  
 CC (W37363), PsbTB (W37363), PsbTC (W37363), PsbTD (W37363),  
 CC (W37363), PsbTE (W37363), PsbTF (W37363), PsbTG (W37363),  
 CC (W37363), PsbTH (W37363), PsbTI (W37363), PsbTJ (W37363),  
 CC (W37363), PsbTK (W37363), PsbTL (W37363), PsbTM (W37363),  
 CC (W37363), PsbTN (W37363), PsbTO (W37363), PsbTP (W37363),  
 CC (W37363), PsbTQ (W37363), PsbTR (W37363), PsbTS (W37363),  
 CC (W37363), PsbTT (W37363), PsbTU (W37363), PsbTV (W37363),  
 CC (W37363), PsbTW (W37363), PsbTX (W37363), PsbTY (W37363),  
 CC (W37363), PsbTZ (W37363), PsbUA (W37363), PsbUB (W37363),  
 CC (W37363), PsbUC (W37363), PsbUD (W37363), PsbUE (W37363),  
 CC (W37363), PsbUF (W37363), PsbUG (W37363), PsbUH (W37363),  
 CC (W37363), PsbUI (W37363), PsbUJ (W37363), PsbUK (W37363),  
 CC (W37363), PsbUL (W37363), PsbUM (W37363), PsbUN (W37363),  
 CC (W37363), PsbUU (W37363), PsbUP (W37363), PsbUQ (W37363),  
 CC (W37363), PsbUR (W37363), PsbUS (W37363), PsbUT (W37363),  
 CC (W37363), PsbUW (W37363), PsbUX (W37363), PsbUY (W37363),  
 CC (W37363), PsbUZ (W37363), PsbVA (W37363), PsbVB (W37363),  
 CC (W37363), PsbVC (W37363), PsbVD (W37363), PsbVE (W37363),  
 CC (W37363), PsbVF (W37363), PsbVG (W37363), PsbVH (W37363),  
 CC (W37363), PsbVI (W37363), PsbVJ (W37363), PsbVK (W37363),  
 CC (W37363), PsbVL (W37363), PsbVM (W37363), PsbVN (W37363),  
 CC (W37363), PsbVO (W37363), PsbVP (W37363), PsbVQ (W37363),  
 CC (W37363), PsbVR (W37363), PsbVS (W37363), PsbVT (W37363),  
 CC (W37363), PsbVW (W37363), PsbVX (W37363), PsbVY (W37363),  
 CC (W37363), PsbVZ (W37363), PsbWA (W37363), PsbWB (W37363),  
 CC (W37363), PsbWC (W37363), PsbWD (W37363), PsbWE (W37363),  
 CC (W37363), PsbWF (W37363), PsbWG (W37363), PsbWH (W37363),  
 CC (W37363), PsbWI (W37363), PsbWJ (W37363), PsbWK (W37363),  
 CC (W37363), PsbWL (W37363), PsbWM (W37363), PsbWN (W37363),  
 CC (W37363), PsbWO (W37363), PsbWP (W37363), PsbWQ (W37363),  
 CC (W37363), PsbWR (W37363), PsbWS (W37363), PsbWT (W37363),  
 CC (W37363), PsbWU (W37363), PsbWV (W37363), PsbWW (W37363),  
 CC (W37363), PsbWX (W37363), PsbWY (W37363), PsbWZ (W37363),  
 CC (W37363), PsbXA (W37363), PsbXB (W37363), PsbXC (W37363),  
 CC (W37363), PsbXD (W37363), PsbXE (W37363), PsbXF (W37363),  
 CC (W37363), PsbXG (W37363), PsbXH (W37363), PsbXI (W37363),  
 CC (W37363), PsbXJ (W37363), PsbXK (W37363), PsbXL (W37363),  
 CC (W37363), PsbXM (W37363), PsbXN (W37363), PsbXO (W37363),  
 CC (W37363), PsbXP (W37363), PsbXQ (W37363), PsbXR (W37363),  
 CC (W37363), PsbXS (W37363), PsbXT (W37363), PsbXU (W37363),  
 CC (W37363), PsbXV (W37363), PsbXW (W37363), PsbXX (W37363),  
 CC (W37363), PsbXY (W37363), PsbXZ (W37363), PsbYA (W37363),  
 CC (W37363), PsbYB (W37363), PsbYC (W37363), PsbYD (W37363),  
 CC (W37363), PsbYE (W37363), PsbYF (W37363), PsbYG (W37363),  
 CC (W37363), PsbYH (W37363), PsbYI (W37363), PsbYJ (W37363),  
 CC (W37363), PsbYK (W37363), PsbYL (W37363), PsbYM (W37363),  
 CC (W37363), PsbYN (W37363), PsbYO (W37363), PsbYP (W37363),  
 CC (W37363), PsbYQ (W37363), PsbYR (W37363), PsbYS (W37363),  
 CC (W37363), PsbYT (W37363), PsbYU (W37363), PsbYV (W37363),  
 CC (W37363), PsbYW (W37363), PsbYX (W37363), PsbYY (W37363),  
 CC (W37363), PsbYZ (W37363), PsbZA (W37363), PsbZB (W37363),  
 CC (W37363), PsbZC (W37363), PsbZD (W37363), PsbZE (W37363),  
 CC (W37363), PsbZF (W37363), PsbZG (W37363), PsbZH (W37363),  
 CC (W37363), PsbZI (W37363), PsbZJ (W37363), PsbZK (W37363),  
 CC (W37363), PsbZL (W37363), PsbZM (W37363), PsbZN (W37363),  
 CC (W37363), PsbZO (W37363), PsbZP (W37363), PsbZQ (W37363),  
 CC (W37363), PsbZR (W37363), PsbZS (W37363), PsbZT (W37363),  
 CC (W37363), PsbZU (W37363), PsbZV (W37363), PsbZW (W37363),  
 CC (W37363), PsbZX (W37363), PsbZY (W37363), PsbZZ (W37363).

Query Match 100.0% Score 39. DB 21. Length 320.  
 Best Local Similarity 100.0% Pred No. 1.69e-02.  
 Matches 5. Conservation 0. Mismatches 0. Indels 0. Gaps 0.

DE 163 WRPP 167  
 |||||  
 QY 1 WRPP 5

RESULT 4  
 ID W09614 standard; Protein; 824 AA.  
 AC W09614  
 DT 20-MAR-1997 (first entry)  
 DE PstA (W09614) protein involved in bacterial synthesis and assembly  
 FW O antigen; pstA; W09614; lipopolysaccharide; infection; diagnosis;  
 KW antibody; dehydratase.  
 OS Staphylococcus aureus strain NCTC 8325-4 (1958).  
 FN W096141974.A1.  
 PD 27-DEC-1996.  
 RF 13-JUN-1996; AU0454.  
 FR 13-JUN-1996; AU0454.  
 CR R-5528; 727221.

PA (AUSN) UNIV AUSTRALIAN NAT.  
 PI Adham El-Adhami W., Matthaei K.L., Stewart P.R.  
 DT 27-MAR-1998 (first entry)  
 DE N-PSDB: T47757.  
 FW Nucleic acids specific to pathogenic Staphylococcus aureus strains -  
 PI for diagnosis and immunogenic protein prodn  
 PS Claim





CC Specific uses are described for each of the polypeptides based on  
CC which tissues they are most highly expressed in, and antibody generating  
CC products for the diagnosis or treatment of cancer, neurodegenerative  
CC disorders, developmental abnormalities and fatal afflictions such as  
CC diseases, tumours, leukemias, diseases of the immune system, autoimmune  
CC diseases, hepatic and renal disease, lymphomas, inflammation, allergies,  
CC ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,  
CC restenosis, prostate diseases, obesity, disorders involving osteoclasts  
CC such as osteoporosis, arthritis or malignancies, diseases of testes, lung  
CC : fibrosis, cystic degeneration disorders, liver cirrhosis and AIDS. The  
CC polypeptides are also useful for identifying their binding partners.  
CC The present sequence represents a hypothetical fragment encoded by a  
CC gene of the invention. (see descriptor line for gene number).  
SQ Sequence 37 AA:

```

Query Match      94 GAG Score 37; DR 38; Length 37;
Best Local Similarity AD AG Pref No 2 gaps(2)
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Lb 32 virpp 36
Gy 1 VVRPP 5
```

PFSULT 9 W8365 Standard Protein: 170 AA.

DE	HUMAN	SECRETED PROTEIN	Human secreted protein fusion protein gene therapy protein therapy;
KW	Human	secreted protein	tumor metastasis disorder leukaemia;
KW	diagnosis	tissue	cancer tumor metastasis disorder blood allergy; renal;
KW	developmental abnormality	foot	deficiency blood allergy; renal;
KW	immune system	ischemia	lymphatic disease brain hepatolymphoma;
KW	inflammation	centamin	dementia Alzheimer's disease restenosis; AIDS;
KW	polymyalgia	fibrositis	myofascitis; fibrositis; osteoarthritis; thyroid; digestion;
KW	osteoporosis	arthritis	testis lung thyroiditis; thyroid; digestion;
KW	endocrine metabolism	regulation	malabsorption gastritis neoplasm;
OS	Homo sapiens		
PN	WO9854963-A2		
PD	10-DEC-1998		
PF	04-JUN-1997	U11422	
PF	18-DEC-1997	US-076923	
PF	06-JUN-1997	US-048877	
PF	06-JUN-1997	US-048881	
PF	06-JUN-1997	US-048884	
PF	06-JUN-1997	US-048893	
PF	06-JUN-1997	US-048896	
PF	06-JUN-1997	US-048899	
PF	06-JUN-1997	US-048915	
PF	06-JUN-1997	US-048949	
PF	06-JUN-1997	US-048964	
PF	06-JUN-1997	US-048972	
PF	06-JUN-1997	US-049020	
PF	06-JUN-1997	US-049375	
PF	05-SEP-1997	US-057628	
PF	05-SEP-1997	US-057635	
PF	05-SEP-1997	US-057644	
PF	05-SEP-1997	US-057647	
PF	05-SEP-1997	US-057650	
PF	05-SEP-1997	US-057661	
PF	05-SEP-1997	US-057667	
PF	05-SEP-1997	US-057761	
PF	05-SEP-1997	US-057764	
PF	05-SEP-1997	US-057770	
PF	05-SEP-1997	US-057775	
PF	05-SEP-1997	US-057778	
PF	06-JUN-1997	US-048875	
PF	06-JUN-1997	US-048878	
PF	06-JUN-1997	US-048882	
PF	06-JUN-1997	US-048887	
PF	06-JUN-1997	US-048894	
PF	06-JUN-1997	US-048900	

PI New isolated human genes and the secreted polypeptides they encode -  
PI useful for diagnosis and treatment of e.g. cancers, neurological  
PI disorders, immune diseases, inflammation or blood disorders  
PS Disclosures, Page 147, 73pp, Enflach.  
CC The invention relates to nucleic acid seq. genes (V84411 to V84633)  
CC encoding human secreted proteins (W82834 to W82876). The secreted protein  
CC gene sequences are deposited with the ATCC under deposit numbers A1CC  
CC 97070, 97071, 97072, 97073, 97074, 97075, 97076, 97077, 97078, 97079, 97080, 97081, 97082, 97083, 97084, 97085, 97086, 97087, 97088, 97089, 97090, 97091, 97092, 97093, 97094, 97095, 97096, 97097, 97098, 97099, 97100, 97101, 97102, 97103, 97104, 97105, 97106, 97107, 97108, 97109, 97110, 97111, 97112, 97113, 97114, 97115, 97116, 97117, 97118, 97119, 97120, 97121, 97122, 97123, 97124, 97125, 97126, 97127, 97128, 97129, 97130, 97131, 97132, 97133, 97134, 97135, 97136, 97137, 97138, 97139, 97140, 97141, 97142, 97143, 97144, 97145, 97146, 97147, 97148, 97149, 97150, 97151, 97152, 97153, 97154, 97155, 97156, 97157, 97158, 97159, 97160, 97161, 97162, 97163, 97164, 97165, 97166, 97167, 97168, 97169, 97170, 97171, 97172, 97173, 97174, 97175, 97176, 97177, 97178, 97179, 97180, 97181, 97182, 97183, 97184, 97185, 97186, 97187, 97188, 97189, 97190, 97191, 97192, 97193, 97194, 97195, 97196, 97197, 97198, 97199, 97200, 97201, 97202, 97203, 97204, 97205, 97206, 97207, 97208, 97209, 97210, 97211, 97212, 97213, 97214, 97215, 97216, 97217, 97218, 97219, 97220, 97221, 97222, 97223, 97224, 97225, 97226, 97227, 97228, 97229, 97230, 97231, 97232, 97233, 97234, 97235, 97236, 97237, 97238, 97239, 97240, 97241, 97242, 97243, 97244, 97245, 97246, 97247, 97248, 97249, 97250, 97251, 97252, 97253, 97254, 97255, 97256, 97257, 97258, 97259, 97260, 97261, 97262, 97263, 97264, 97265, 97266, 97267, 97268, 97269, 97270, 97271, 97272, 97273, 97274, 97275, 97276, 97277, 97278, 97279, 97280, 97281, 97282, 97283, 97284, 97285, 97286, 97287, 97288, 97289, 97290, 97291, 97292, 97293, 97294, 97295, 97296, 97297, 97298, 97299, 97300, 97301, 97302, 97303, 97304, 97305, 97306, 97307, 97308, 97309, 97310, 97311, 97312, 97313, 97314, 97315, 97316, 97317, 97318, 97319, 9732





11  
 12 100% standard: Protein: 1157 AA.  
 13 100%  
 14 11 100% (first entry)  
 15 12 100%  
 16 13 100%  
 17 14 100%  
 18 15 100%  
 19 16 100%  
 20 17 100%  
 21 18 100%  
 22 19 100%  
 23 20 100%  
 24 21 100%  
 25 22 100%  
 26 23 100%  
 27 24 100%  
 28 25 100%  
 29 26 100%  
 30 27 100%  
 31 28 100%  
 32 29 100%  
 33 30 100%  
 34 31 100%  
 35 32 100%  
 36 33 100%  
 37 34 100%  
 38 35 100%  
 39 36 100%  
 40 37 100%  
 41 38 100%  
 42 39 100%  
 43 40 100%  
 44 41 100%  
 45 42 100%  
 46 43 100%  
 47 44 100%  
 48 45 100%  
 49 46 100%  
 50 47 100%  
 51 48 100%  
 52 49 100%  
 53 50 100%  
 54 51 100%  
 55 52 100%  
 56 53 100%  
 57 54 100%  
 58 55 100%  
 59 56 100%  
 60 57 100%  
 61 58 100%  
 62 59 100%  
 63 60 100%  
 64 61 100%  
 65 62 100%  
 66 63 100%  
 67 64 100%  
 68 65 100%  
 69 66 100%  
 70 67 100%  
 71 68 100%  
 72 69 100%  
 73 70 100%  
 74 71 100%  
 75 72 100%  
 76 73 100%  
 77 74 100%  
 78 75 100%  
 79 76 100%  
 80 77 100%  
 81 78 100%  
 82 79 100%  
 83 80 100%  
 84 81 100%  
 85 82 100%  
 86 83 100%  
 87 84 100%  
 88 85 100%  
 89 86 100%  
 90 87 100%  
 91 88 100%  
 92 89 100%  
 93 90 100%  
 94 91 100%  
 95 92 100%  
 96 93 100%  
 97 94 100%  
 98 95 100%  
 99 96 100%  
 100 97 100%  
 101 98 100%  
 102 99 100%  
 103 100 100%  
 104 101 100%  
 105 102 100%  
 106 103 100%  
 107 104 100%  
 108 105 100%  
 109 106 100%  
 110 107 100%  
 111 108 100%  
 112 109 100%  
 113 110 100%  
 114 111 100%  
 115 112 100%  
 116 113 100%  
 117 114 100%  
 118 115 100%  
 119 116 100%  
 120 117 100%  
 121 118 100%  
 122 119 100%  
 123 120 100%  
 124 121 100%  
 125 122 100%  
 126 123 100%  
 127 124 100%  
 128 125 100%  
 129 126 100%  
 130 127 100%  
 131 128 100%  
 132 129 100%  
 133 130 100%  
 134 131 100%  
 135 132 100%  
 136 133 100%  
 137 134 100%  
 138 135 100%  
 139 136 100%  
 140 137 100%  
 141 138 100%  
 142 139 100%  
 143 140 100%  
 144 141 100%  
 145 142 100%  
 146 143 100%  
 147 144 100%  
 148 145 100%  
 149 146 100%  
 150 147 100%  
 151 148 100%  
 152 149 100%  
 153 150 100%  
 154 151 100%  
 155 152 100%  
 156 153 100%  
 157 154 100%  
 158 155 100%  
 159 156 100%  
 160 157 100%  
 161 158 100%  
 162 159 100%  
 163 160 100%  
 164 161 100%  
 165 162 100%  
 166 163 100%  
 167 164 100%  
 168 165 100%  
 169 166 100%  
 170 167 100%  
 171 168 100%  
 172 169 100%  
 173 170 100%  
 174 171 100%  
 175 172 100%  
 176 173 100%  
 177 174 100%  
 178 175 100%  
 179 176 100%  
 180 177 100%  
 181 178 100%  
 182 179 100%  
 183 180 100%  
 184 181 100%  
 185 182 100%  
 186 183 100%  
 187 184 100%  
 188 185 100%  
 189 186 100%  
 190 187 100%  
 191 188 100%  
 192 189 100%  
 193 190 100%  
 194 191 100%  
 195 192 100%  
 196 193 100%  
 197 194 100%  
 198 195 100%  
 199 196 100%  
 200 197 100%  
 201 198 100%  
 202 199 100%  
 203 200 100%  
 204 201 100%  
 205 202 100%  
 206 203 100%  
 207 204 100%  
 208 205 100%  
 209 206 100%  
 210 207 100%  
 211 208 100%  
 212 209 100%  
 213 210 100%  
 214 211 100%  
 215 212 100%  
 216 213 100%  
 217 214 100%  
 218 215 100%  
 219 216 100%  
 220 217 100%  
 221 218 100%  
 222 219 100%  
 223 220 100%  
 224 221 100%  
 225 222 100%  
 226 223 100%  
 227 224 100%  
 228 225 100%  
 229 226 100%  
 230 227 100%  
 231 228 100%  
 232 229 100%  
 233 230 100%  
 234 231 100%  
 235 232 100%  
 236 233 100%  
 237 234 100%  
 238 235 100%  
 239 236 100%  
 240 237 100%  
 241 238 100%  
 242 239 100%  
 243 240 100%  
 244 241 100%  
 245 242 100%  
 246 243 100%  
 247 244 100%  
 248 245 100%  
 249 246 100%  
 250 247 100%  
 251 248 100%  
 252 249 100%  
 253 250 100%  
 254 251 100%  
 255 252 100%  
 256 253 100%  
 257 254 100%  
 258 255 100%  
 259 256 100%  
 260 257 100%  
 261 258 100%  
 262 259 100%  
 263 260 100%  
 264 261 100%  
 265 262 100%  
 266 263 100%  
 267 264 100%  
 268 265 100%  
 269 266 100%  
 270 267 100%  
 271 268 100%  
 272 269 100%  
 273 270 100%  
 274 271 100%  
 275 272 100%  
 276 273 100%  
 277 274 100%  
 278 275 100%  
 279 276 100%  
 280 277 100%  
 281 278 100%  
 282 279 100%  
 283 280 100%  
 284 281 100%  
 285 282 100%  
 286 283 100%  
 287 284 100%  
 288 285 100%  
 289 286 100%  
 290 287 100%  
 291 288 100%  
 292 289 100%  
 293 290 100%  
 294 291 100%  
 295 292 100%  
 296 293 100%  
 297 294 100%  
 298 295 100%  
 299 296 100%  
 300 297 100%  
 301 298 100%  
 302 299 100%  
 303 300 100%  
 304 301 100%  
 305 302 100%  
 306 303 100%  
 307 304 100%  
 308 305 100%  
 309 306 100%  
 310 307 100%  
 311 308 100%  
 312 309 100%  
 313 310 100%  
 314 311 100%  
 315 312 100%  
 316 313 100%  
 317 314 100%  
 318 315 100%  
 319 316 100%  
 320 317 100%  
 321 318 100%  
 322 319 100%  
 323 320 100%  
 324 321 100%  
 325 322 100%  
 326 323 100%  
 327 324 100%  
 328 325 100%  
 329 326 100%  
 330 327 100%  
 331 328 100%  
 332 329 100%  
 333 330 100%  
 334 331 100%  
 335 332 100%  
 336 333 100%  
 337 334 100%  
 338 335 100%  
 339 336 100%  
 340 337 100%  
 341 338 100%  
 342 339 100%  
 343 340 100%  
 344 341 100%  
 345 342 100%  
 346 343 100%  
 347 344 100%  
 348 345 100%  
 349 346 100%  
 350 347 100%  
 351 348 100%  
 352 349 100%  
 353 350 100%  
 354 351 100%  
 355 352 100%  
 356 353 100%  
 357 354 100%  
 358 355 100%  
 359 356 100%  
 360 357 100%  
 361 358 100%  
 362 359 100%  
 363 360 100%  
 364 361 100%  
 365 362 100%  
 366 363 100%  
 367 364 100%  
 368 365 100%  
 369 366 100%  
 370 367 100%  
 371 368 100%  
 372 369 100%  
 373 370 100%  
 374 371 100%  
 375 372 100%  
 376 373 100%  
 377 374 100%  
 378 375 100%  
 379 376 100%  
 380 377 100%  
 381 378 100%  
 382 379 100%  
 383 380 100%  
 384 381 100%  
 385 382 100%  
 386 383 100%  
 387 384 100%  
 388 385 100%  
 389 386 100%  
 390 387 100%  
 391 388 100%  
 392 389 100%  
 393 390 100%  
 394 391 100%  
 395 392 100%  
 396 393 100%  
 397 394 100%  
 398 395 100%  
 399 396 100%  
 400 397 100%  
 401 398 100%  
 402 399 100%  
 403 400 100%  
 404 401 100%  
 405 402 100%  
 406 403 100%  
 407 404 100%  
 408 405 100%  
 409 406 100%  
 410 407 100%  
 411 408 100%  
 412 409 100%  
 413 410 100%  
 414 411 100%  
 415 412 100%  
 416 413 100%  
 417 414 100%  
 418 415 100%  
 419 416 100%  
 420 417 100%  
 421 418 100%  
 422 419 100%  
 423 420 100%  
 424 421 100%  
 425 422 100%  
 426 423 100%  
 427 424 100%  
 428 425 100%  
 429 426 100%  
 430 427 100%  
 431 428 100%  
 432 429 100%  
 433 430 100%  
 434 431 100%  
 435 432 100%  
 436 433 100%  
 437 434 100%  
 438 435 100%  
 439 436 100%  
 440 437 100%  
 441 438 100%  
 442 439 100%  
 443 440 100%  
 444 441 100%  
 445 442 100%  
 446 443 100%  
 447 444 100%  
 448 445 100%  
 449 446 100%  
 450 447 100%  
 451 448 100%  
 452 449 100%  
 453 450 100%  
 454 451 100%  
 455 452 100%  
 456 453 100%  
 457 454 100%  
 458 455 100%  
 459 456 100%  
 460 457 100%  
 461 458 100%  
 462 459 100%  
 463 460 100%  
 464 461 100%  
 465 462 100%  
 466 463 100%  
 467 464 100%  
 468 465 100%  
 469 466 100%  
 470 467 100%  
 471 468 100%  
 472 469 100%  
 473 470 100%  
 474 471 100%  
 475 472 100%  
 476 473 100%  
 477 474 100%  
 478 475 100%  
 479 476 100%  
 480 477 100%  
 481 478 100%  
 482 479 100%  
 483 480 100%  
 484 481 100%  
 485 482 100%  
 486 483 100%  
 487 484 100%  
 488 485 100%  
 489 486 100%  
 490 487 100%  
 491 488 100%  
 492 489 100%  
 493 490 100%  
 494 491 100%  
 495 492 100%  
 496 493 100%  
 497 494 100%  
 498 495 100%  
 499 496 100%  
 500 497 100%  
 501 498 100%  
 502 499 100%  
 503 500 100%  
 504 501 100%  
 505 502 100%  
 506 503 100%  
 507 504 100%  
 508 505 100%  
 509 506 100%  
 510 507 100%  
 511 508 100%  
 512 509 100%  
 513 510 100%  
 514 511 100%  
 515 512 100%  
 516 513 100%  
 517 514 100%  
 518 515 100%  
 519 516 100%  
 520 517 100%  
 521 518 100%  
 522 519 100%  
 523 520 100%  
 524 521 100%  
 525 522 100%  
 526 523 100%  
 527 524 100%  
 528 525 100%  
 529 526 100%  
 530 527 100%  
 531 528 100%  
 532 529 100%  
 533 530 100%  
 534 531 100%  
 535 532 100%  
 536 533 100%  
 537 534 100%  
 538 535 100%  
 539 536 100%  
 540 537 100%  
 541 538 100%  
 542 539 100%  
 543 540 100%  
 544 541 100%  
 545 542 100%  
 546 543 100%  
 547 544 100%  
 548 545 100%  
 549 546 100%  
 550 547 100%  
 551 548 100%  
 552 549 100%  
 553 550 100%  
 554 551 100%  
 555 552 100%  
 556 553 100%  
 557 554 100%  
 558 555 100%  
 559 556 100%  
 560 557 100%  
 561 558 100%  
 562 559 100%  
 563 560 100%  
 564 561 100%  
 565 562 100%  
 566 563 100%  
 567 564 100%  
 568 565 100%  
 569 566 100%  
 570 567 100%  
 571 568 100%  
 572 569 100%  
 573 570 100%  
 574 571 100%  
 575 572 100%  
 576 573 100%  
 577 574 100%  
 578 575 100%  
 579 576 100%  
 580 577 100%  
 581 578 100%  
 582 579 100%  
 583 580 100%  
 584 581 100%  
 585 582 100%  
 586 583 100%  
 587 584 100%  
 588 585 100%  
 589 586 100%  
 590 587 100%  
 591 588 100%  
 592 589 100%  
 593 590 100%  
 594 591 100%  
 595 592 100%  
 596 593 100%  
 597 594 100%  
 598 595 100%  
 599 596 100%  
 600 597 100%  
 601 598 100%  
 602 599 100%  
 603 600 100%  
 604 601 100%  
 605 602 100%  
 606 603 100%  
 607 604 100%  
 608 605 100%  
 609 606 100%  
 610 607 100%  
 611 608 100%  
 612 609 100%  
 613 610 100%  
 614 611 100%  
 615 612 100%  
 616 613 100%  
 617 614 100%  
 618 615 100%  
 619 616 100%  
 620 617 100%  
 621 618 100%  
 622 619 100%  
 623 620 100%  
 624 621 100%  
 625 622 100%  
 626 623 100%  
 627 624 100%  
 628 625 100%  
 629 626 100%  
 630 627 100%  
 631 628 100%  
 632 629 100%  
 633 630 100%  
 634 631 100%  
 635 632 100%  
 636 633 100%  
 637 634 100%  
 638 635 100%  
 639 636 100%  
 640 637 100%  
 641 638 100%  
 642 639 100%  
 643 640 100%  
 644 641 100%  
 645 642 100%  
 646 643 100%  
 647 644 100%  
 648 645 100%  
 649 646 100%  
 650 647 100%  
 651 648 100%  
 652 649 100%  
 653 650 100%  
 654 651 100%  
 655 652 100%  
 656 653 100%  
 657 654 100%  
 658 655 100%  
 659 656 100%  
 660 657 100%  
 661 658 100%  
 662 659 100%  
 663 660 100%  
 664 661 100%  
 665 662 100%  
 666 663 100%  
 667 664 100%  
 668 665 100%  
 669 666 100%  
 670 667 100%  
 671 668 100%  
 672 669 100%  
 673 670 100%  
 674 671 100%  
 675 672 100%  
 676 673 100%  
 677 674 100%  
 678 675 100%  
 679 676 100%  
 680 677 100%  
 681 678 100%  
 682 679 100%  
 683 680 100%  
 684 681 100%  
 685 682 100%  
 686 683 100%  
 687 684 100%  
 688 685 100%  
 689 686 100%  
 690 687 100%  
 691 688 100%  
 692 689 100%  
 693 690 100%  
 694 691 100%  
 695 692 100%  
 696 693 100%  
 697 694 100%  
 698 695 100%  
 699











Query Match 94.4%; Score 47.14; length 1227;  
 Best Local Similarity 80.0%; Pred. No. 2.17e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Db 325 VWRPP 329  
 QY 1 VWRPP 5

Search completed. Fri Feb 25 13:19:44 2000  
 Job time : 45 secs.

\*\*\*\*\*  
[WATERFALL] (IM)  
\*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit,  
Copyright (c) 1994-1999 University of Pittsburgh, PA  
Distribution rights by Oxford Molecular Ltd

Misfold\_PP Protein - protein database search using Smith-Waterman algorithm  
Score: 68.411 E-164 25 13.15-24 2153. Masfar time 2.33 seconds  
Tabular output not generated.

Database: US-09-132-799-2  
(15) from US95182799.ppt

Searcher: 1 VVRPP 5

Searcher: FAM 150  
Cap 15

Searcher: 122820 seqs, 4006643 residues

Post-processor: Minimum Match 90%

Database: 1000 summaries

Database: 1000 summaries

Database: 1000 summaries

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description	Prod. No.
1	86	100.0	172	2 S280R2	vif protein - simian	1.73e-01
2	86	100.0	183	2 S28220	cruciferin (c10a) PAG	1.73e-01
3	86	100.0	196	2 S28218	cruciferin (clone PAG	1.73e-01
4	86	100.0	211	2 S28162	3SK prolactin-rich prot	1.73e-01
5	86	100.0	212	2 S07652	coat protein - pea ro	1.73e-01
6	86	100.0	234	2 S28217	cruciferin (c10a) PAG	1.73e-01
7	86	100.0	235	2 S28260	proline-rich cell wall	1.73e-01
8	86	100.0	242	2 A48825	hypothetical 32K prot	1.73e-01
9	86	100.0	425	2 A47148	beta-lactamase (EC 3.	1.73e-01
10	86	100.0	481	2 S52985	cell wall protein - a	1.73e-01
11	86	100.0	474	2 S04565	legumin precursor (cl	1.73e-01
12	86	100.0	476	2 E14119	probable indole-3-ace	1.73e-01
13	86	100.0	479	2 S28223	cruciferin precursor,	1.73e-01
14	86	100.0	481	2 T01850	UDP-glucose glucosylt	1.73e-01
15	86	100.0	509	2 A45540	cruciferin 1 precursor	1.73e-01
16	86	100.0	534	2 S41735	cholesterol esterase	1.73e-01
17	86	100.0	538	2 S05684	lipase 1 precursor -	1.73e-01
18	86	100.0	549	2 S05553	triacylglycerol lipas	1.73e-01
19	86	100.0	549	2 S24448	triacylglycerol lipas	1.73e-01
20	86	100.0	549	2 JN0551	triacylglycerol lipas	1.73e-01
21	86	100.0	706	2 S62933	hypothetical protein	1.73e-01
22	86	100.0	773	2 A47666	structural polyprotei	1.73e-01
23	86	100.0	1224	2 S41199	probable adenylate/4u	1.73e-01

24	39	106.0	186	1 GNWVIF	decapne protein (c10a)	1.7e-01
25	37	94.9	99	2 S29903	modulation protein ho	5.0e-01
26	37	94.9	183	2 S20844	modulation protein ho	5.0e-01
27	37	94.9	230	2 S07100	probable serine prote	5.0e-01
28	37	94.9	233	2 H7666	cruciferin (c10a) PAG	5.0e-01
29	37	94.9	233	2 S28223	cruciferin (c10a) PAG	5.0e-01
30	37	94.9	268	2 S71427	conserved hypothetical	5.0e-01
31	37	94.9	277	2 S5355	conserved hypothetical	5.0e-01
32	37	94.9	282	2 S04663	cell wall protein - a	5.0e-01
33	37	94.9	341	2 S41111	cruciferin precursor	5.0e-01
34	37	94.9	371	2 S51320	mitochondrial lipid pr	5.0e-01
35	37	94.9	484	2 S66233	N-ethylmaleimide etha	5.0e-01
36	37	94.9	492	2 C71445	hypothetical protein	5.0e-01
37	37	94.9	440	2 S4968	glyoxaldehyde 3-thus	5.0e-01
38	37	94.9	450	2 S41111	probable serine prote	5.0e-01
39	37	94.9	455	1 S04510	cruciferin precursor	5.0e-01
40	37	94.9	465	1 S14742	cruciferin 4 precursor	5.0e-01
41	37	94.9	472	2 S08509	cruciferin precursor	5.0e-01
42	37	94.9	484	2 S71419	probable indole-3-ace	5.0e-01
43	37	94.9	486	2 S37543	cruciferin 1 precursor	5.0e-01
44	37	94.9	493	2 S40889	11S albumin seed stro	5.0e-01
45	37	94.9	496	2 S25091	cruciferin R02 (cap	5.0e-01
46	37	94.9	548	2 S42615	triacylglycerol lipas	5.0e-01
47	37	94.9	627	2 A40201	artificial waxing soap	5.0e-01
48	37	94.9	810	2 S50849	PKK1 protein - yeast	5.0e-01
49	37	94.9	822	2 S38728	epidermal growth fact	5.0e-01
50	37	94.9	940	2 A40646	proectin - fruit fly	5.0e-01
51	37	94.9	1031	1 A38713	kinasin heavy chain	5.0e-01
52	37	94.9	1242	2 S14251	probable 3-ketoacyl-CoA	5.0e-01
53	37	94.9	1265	2 S57668	benkinding protein 2	5.0e-01

Note: Post-processor removed 947 summaries from list due to search parameters chosen.

ALIGNMENTS

RESULT	1					
ENTRY	S280R2	***type complete				
TITLE	vif protein - simian immunodeficiency virus					
ORGANISM	Simian immunodeficiency virus, HIV					
DATE	17-Apr-1993	Sequence Division 17-Apr-1993	0.02	0.02	0.02	0.02
ADDITIONS	S280R2	17-Mar-1994				
REFERENCE	S280R2					
AUTHORS	Isajima, H.; Hasegawa, A.; Maki, N.; Fukasawa, M.; Mizuno, T.; Spiegel, S.; Cooper, K.W.; Miyama, E.N.; Goto, T.; Hayami, M.					
JOURNAL	Nature (1989) 341:539-541					
TITLE	Sequence of a novel simian immunodeficiency virus from a wild-caught African mandrill.					
CROSS-REFERENCES	M010:90015168					
STATUS	S280R2					
STATUS	***status	nucleic acid sequence not shown; translation not shown				
PROTEIN	172	***label 189				
RESIDUES	172	***label 189				
SEQUENCE	the nucleotide sequence was submitted to the EMBL Data Library, July 1989					
GENETICS						
GENE	vif					
SUMMARY	172	***molecular-weight: 20400				
Query Match	100.0%	Score 36	DB 21	Length 172		
Best Local Similarity	100.0%	Prod. No. 1.74e-01				
Matches	5	Conservative 5	Mismatches 1	Indels 0		
DB	95 VVRPP 99					
QV	1 VVRPP 5					
RESULT	2					
ENTRY	S28220	***type fragment				

```

TITLE      cruciferin (clone pAG1) - radish (fragment)
ALTERNATE_NAMES  seed storage protein
ORGANISM    "Formal name Raphanus sativus #common name radish"
DATE        11-Mar-1998 #sequence_revision 17-Apr 1998 #text_change
05-Jun-1998
ACCESSIONS  S24200
REFERENCE    Grellier, F., Delaunay-Thomas, E., Paynal, M., Aspart, L., Delsonoy, M.,
            Grellier, F., Pail, (1998) 20,477-479
            #journal
            #title      The cruciferin gene family in radish.
            #cross-references EMBL:X59803
            #accession  S24200
            #molecule_type mRNA
            #residues    1196 #label DEP
            #cross-references EMBL:X59803
            #experimental_source clone pAG1
            #classification #superfamily glycinin
            #keywords     seed, storage protein
            #length 183 #checksum 7244

Query Match 100.0% Score 39 DP 2 Length 183
Best Local Similarity 100.0% Pred. No. 1,73e-01
Matches 5: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 174 VWRPP 178
      |||||
QY 1 VWRPP 5

RESULT 3
ENTRY  #type fragment
TITLE  cruciferin (clone pAG1) - radish (fragment)
ALTERNATE_NAMES  seed storage protein
ORGANISM    "Formal name Raphanus sativus #common name radish"
DATE        11-Mar-1998 #sequence_revision 17-Apr 1998 #text_change
05-Jun-1998
ACCESSIONS  S24218, S24244
REFERENCE    Grellier, F., Delaunay-Thomas, E., Paynal, M., Aspart, L., Delsonoy, M.,
            Grellier, F., Pail, (1998) 20,477-479
            #journal
            #title      The cruciferin gene family in radish.
            #cross-references EMBL:X59803
            #accession  S24218, S24244
            #molecule_type mRNA
            #residues    1196 #label DEP
            #cross-references EMBL:X59803
            #experimental_source clone pAG1
            #classification #superfamily glycinin
            #keywords     seed, storage protein
            #length 183 #checksum 9731

Query Match 100.0% Score 39 DP 2 Length 183
Best Local Similarity 100.0% Pred. No. 1,73e-01
Matches 5: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 11 VWRPP 15
      |||||
QY 1 VWRPP 5

RESULT 4
ENTRY  #type fragment
TITLE  33K proline-rich protein - carrot (fragment)
ORGANISM    "Formal name Daucus carota #common name carrot"

```

```

DATE        08-Aug-1987 #sequence_revision 08-Aug-1987 #text_change
30-Sep-1993
ACCESSIONS  B22162
REFERENCE    A94061
            Chen, J., Varner, J.E.,
            Ford, N.H., Azai, S.H., U.S.A. (1995) 40,4399-4403
            #journal
            #title      Isolation and characterization of cDNA clones for carrot
            #accession  B22162
            #molecule_type mRNA
            #residues    1211 #label CHE
            #classification #superfamily proline-rich protein 3
            #summary      #length 211 #checksum 2886

Query Match 100.0% Score 39 DP 2 Length 211
Best Local Similarity 100.0% Pred. No. 1,73e-01
Matches 5: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 184 VWRPP 188
      |||||
QY 1 VWRPP 5

RESULT 5
ENTRY  #type complete
TITLE  coat protein - pea early browning virus
ORGANISM    "Formal name pea early browning virus"
VARIETY      strain SP5
DATE        07-Sep-1996 #sequence_revision 07-Sep-1996 #text_change
08-Sep-1997
ACCESSIONS  S07562, S14568, S08695
REFERENCE    Peterson, S.G., Lehmbeck, J., Berthardt, B.,
            Plant, M., Piel, (1996) 12,705-737
            #journal
            #title      Analysis of RNAi of pea early browning virus strain SP5.
            #cross-references EMBL:X59803
            #accession  S07562
            #molecule_type genomic RNA
            #residues    1212 #label PET
            #cross-references EMBL:X59803
            #accession  S14568
            #journal
            #title      Grellier, M.G., Delaunay-Thomas, E., Paynal, M., Aspart, L., Pail,
            #cross-references EMBL:X59803
            #accession  S14568
            #molecule_type genomic RNA
            #residues    1212 #label GOM
            #cross-references EMBL:X59803
            #experimental_source British strain SP5 isolate

GENETICS
            #map_position segment 2
            #classification #superfamily coat protein
            #summary      #length 212 #checksum 6766

Query Match 100.0% Score 39 DP 2 Length 212
Best Local Similarity 100.0% Pred. No. 1,73e-01
Matches 5: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 127 VWRPP 131
      |||||
QY 1 VWRPP 5

RESULT 6
ENTRY  #type fragment
TITLE  cruciferin (clone pAF7) - radish (fragment)
ALTERNATE_NAMES  seed storage protein
ORGANISM    "Formal name Raphanus sativus #common name radish"
DATE        11-Mar-1998 #sequence_revision 17-Apr 1998 #text_change

```



```

##cross-references EMBL:M36761
##experimental-source strain Rcc S332, clone pM51
##note the authors designated the arg codon for residue 1 as
##note start codon; the GTC codon for residue 1 was
translated as Val
#accession S27006
#molecule-type protein
#residues 47-57189 96 #label 1B2
#experimental-source overexpressed cloned gene from strain F03-S-152
REFERENCE S01458
#authors de Weester, P.; Joris, B.; Lenzini, M.V.; Delhotay, P.;
Ercilim, T.; Dussart, P.; Klein, D.; Ohysen, M.; Provo,
J.M.; van Beumen, J.
#journal Biochem. J. (198) 244:427-432
#title The active sites of the beta-lactamase of Streptomyces
cacaoli and Streptomyces albus 3.
#cross-references MIM:581411
#accession S11762
#molecule-type protein
#residues 89-96 #label DEM
#experimental-source cloned gene from plasmid pMX51 overexpressed in
Streptomyces lividans
COMMENT This is a class A beta-lactamase.
CLASSIFICATION #superfamily beta-lactamase I
KEYWORDS antibiotic resistance; extracellular protein; hydrolase
FEATURE
1-45 #domain signal sequence #status predicted #label SIG
47-238 #product beta-lactamase #status experimental #label MAT
93 #active-site Ser #status experimental
SUMMARY #length 325 #molecular-weight 37557 #checksum 2127
Query Match 100.0% Score 32.18 E-16 Length 325
Best Local Similarity 100.0% Pred No. 1.73e-016
Matches 5 Conservative 0 Mismatches 0 Indels 0 Gaps 0
QY 1 VRPP 5
EL 291 VRPP 238
|||||
QY 1 VRPP 5

RESULT 10
ENTRY S52985 #type complete
TITLE Cell wall protein - alfalfa
#formal_name Medicago sativa #common_name alfalfa
ORGANISM 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change
DATE 04-Sep-1998
ACCESSIONS S52985
REFERENCE S52985
#authors Neufch, C.E.; Winicov, I.
#journal Plant Mol. Biol. (1995) 27:411-418
#title Post-transcriptional regulation of a salt-inducible alfalfa
gene encoding a putative plasma membrane protein
#cross-references MIM:5619167
#accession S52985
#status preliminary
#molecule-type RNA
#residues 1-381 #label pM
#cross-references EMBL:U0017 NID_05596120 F00056913
CLASSIFICATION #superfamily beta-lactamase
SUMMARY #length 392 #molecular-weight 42620 #checksum 4221
Query Match 100.0% Score 39.18 E-2 Length 391
Best Local Similarity 100.0% Pred No. 1.73e-017
Matches 5 Conservative 0 Mismatches 0 Indels 0 Gaps 0
QY 1 VRPP 5
Db 146 VRPP 150
|||||
QY 1 VRPP 5

RESULT 11
ENTRY S68965 #type complete
TITLE
#formal_name Magnolia salicifolia
#journal 14-Mar-1997 #sequence_revision 14-Mar-1997 #text_change
DATE 04-Sep-1997
ACCESSIONS S68965
REFERENCE S68965
#authors Fischer, H.; Haake, V.; Horstmann, C.; Jensen, U.;
Eck, J.; Blochom, (1995) 229:645-650
#journal Chemotaxonomy and molecular phylogeny: relationships of Magnolia
legumin-encoding cDNAs representing two divergent gene
subfamilies.
#cross-references MIM:0929015
#accession S68965
#molecule-type mRNA
#residues 1-474 #label FIS1
#cross-references MIM: N93464 MIM:0929015 FIS 273256
#note the authors translated the cDNA 273 for residue 450 as
Val
#accession S77067
#molecule-type protein
#residues 49-105 #label FIS2
CLASSIFICATION #superfamily glycinin
KEYWORDS seed; storage protein
FEATURE
1-23 #domain signal sequence #status predicted #label SIG
24-299 #product legumin alpha chain #status experimental #label
MAT1
SUMMARY #length 475 #molecular-weight 53000 #checksum 7790
Query Match 100.0% Score 43.18 E-16 Length 475
Best Local Similarity 100.0% Pred No. 1.73e-016
Matches 5 Conservative 0 Mismatches 0 Indels 0 Gaps 0
QY 1 VRPP 5
EL 270 VRPP 274
|||||
QY 1 VRPP 5

RESULT 12
ENTRY E71419 #type complete
TITLE Probable indole-3-acetate beta-glucosyltransferase -
Arabidopsis thaliana
#formal_name Arabidopsis thaliana #common_name mouse-ear
cress
ORGANISM Columbia
#variety Columbia
DATE 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change
ACCESSIONS E71419
REFERENCE E71419
#authors Pao, M.; Pao, F.; Pao, F.; Pao, F.; Pao, F.; Pao, F.;
Dean, C.; Bergkamp, R.; Dirkse, W.; Van Staveren, M.;
Stiekema, W.; Frost, L.; Ridley, P.; Hudson, S.A.; Patel,
P.; Murphy, S.; Piffanelli, P.; Wedler, H.; Wedler, E.;
Wambutt, R.; Wittenberger, R.; Pohl, T.; Terry, M.;
Gleason, J.; Villarroel, R.; De Clerck, R.; Van Montagu, M.;
Lechardy, A.; Aubert, S.; Gy, I.; Kreis, M.; Lao, N.;
Kavanagh, T.; Hempel, S.; Kottler, P.; Entian, K.D.; Rieger,
M.; Schaefer, M.; Fink, B.; Meller, A.; S. Silver, M.;
Tamas, P.; Maffei, A.; Pao, F.; Pao, F.; Pao, F.; Pao, F.;
A. Voulgaris, E.; Miloudi, D.; Hatzopoulos, P.;
Piravandi, E.; Obermaier, B.; Hilbert, H.; Duesterhoft, A.;
Morges, T.; Jones, T.D.; Fueva, T.; Palmer, K.; Benes, V.;
Kutmar, S.; Acosta, W.; Cacko, P.; Berger, C.; Delsen,
M.; Voet, M.; Volckaert, G.; Mewes, H.W.; Klosterman, S.;
Schueller, C.; Chalwatzis, N.
#journal Nature (1998) 391:485-488
#title Analysis of 1.4 Mb of contiguous sequence from chromosome 4
of Arabidopsis thaliana.
#cross-references MIM:08121113
#accession E71419
#status preliminary; nucleic acid sequence not shown;

```



Query Match 100.0%; Score 39; DB 2; Length 509;  
 Best Local Similarity 100.0%; Pred No 1 73e-01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 292 VRRP 296  
 |||||  
 Qy 1 VRRP 5

Search completed: Fri Feb 25 13:14:13 2000  
 Job time : 49 secs.





```

DE PROLINE RICH 35 KD EXTENSIN-RELATED PROTEIN PREPARED FOR FRAGMENT
OS LADUS CAROTA (CARROT)
OC EUCARYOTA: VIRIDIPHYTES: STEPTOPHYTES: EMERYOPHYTES: FRAGIDOPHYTES:
OC EUBYLICONTES: SPERMATOPHYTES: MAGNOLIOPHYTES: EUPHYTIOPHYTES:
OC ASTERIDAE: ARABIDAE: APICAE: DACTYL.
RN [1]
RP SEQUENCE FROM N.A.
RA CHEN J., VARNER J.E.,
RI "Isolation and characterization of cDNA clones for carrot extensin
PI PROX NALL ACAP SCI U.S.A. 82:4399-4403(1995)
CI "SIMILARITY: BELONGS TO THE EMERIC FAMILY OF EXTENSINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.ist-sib.ch/announcements
CC or send an email to license@sib-sib.ch.)
CC -----
DE EMBL: M1222; GI: 67652;
DE PIR: P2162; POSITIVE
KW CELL WALL; REPEAT; SIGNAL
FT NON-TER
FT SIGNAL <1 14
FT CHAIN 15 211
FT PROLINE RICH 35 KD EXTENSIN-RELATED
FT VARIAT 29 29
FT VARIANT 22 22
FT SEQUENCE 211 AA: 21621 MW: 540266P QRC32:
Query Match 100.0% Score 29 DB 1 Length 211
Rest Local Similarity 100.0% Prid No 5 466-00
Matches 5 Conservative 0 Mismatches 0 Indels 0 Gaps 0
Db 184 VWRPP 188
11111
QY 1 VWRPP 5
RESULT 3
ID COAT-PEEV STANDARD: PRT: 212 AA.
AC P14849:
DI 01-APR-1990 (REL 14, CREATED)
DI 01-APR-1990 (REL 14, LAST SEQUENCE UPDATE)
DI 01-MAR-1992 (REL 23, LAST ANNOTATION UPDATE)
DE COAT PROTEIN (CAPSID PROTEIN)
OS PEA EARLY BROWNING VIRUS
OC VIRUSES: GRESSA POSITIVE-STRAND: VIRUSES, NO DNA STABLE, TORRAVIRUS
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S95:
RA MEDLINE: 91370950.
RA PETERSEN S.G., IFFENHART J., ROPPHART P.:
RI "Analysis of the RNA of pea early browning virus strain S95."
RI PLANT MOL. BIOL. 13:735-737(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S95:
RA MEDLINE: 90356386.
RA GOURDON M.G., LOMONOSOFF G.P., DAVIES J.W., WOOD K.R.:
RI "The complete nucleotide sequence of PEBV RNA: reveals the presence
PI of a novel open reading frame and provides insights into the
PI structure of toroviral subgenomic promoters."
RI NUCLEIC ACIDS RES. 18:4507-4512(1990).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.ist-sib.ch/announcements
CC or send an email to license@sib-sib.ch.)

```

```

CC or send an email to license@sib-sib.ch.)
CC -----
DE EMBL: X1688; GI: 62041;
DE PIR: N01928; Q00067;
DE PIR: S07552; S07553.
DE PIR: S14698; S14698.
KW COAT PROTEIN
SI SEQUENCE 212 AA: 20244 MW: 11003999 QRC32:
Query Match 100.0% Score 39 DB 1 Length 212
Rest Local Similarity 100.0% Prid No 5 466-00
Matches 5 Conservative 0 Mismatches 0 Indels 0 Gaps 0
Db 161 VWRPP 165
11111
QY 1 VWRPP 5
RESULT 5
ID RAL-STRTI STANDARD: PRT: 325 AA.
AC Q03680:
DI 01-FEB-1994 (REL 28, CREATED)
DI 01-FEB-1994 (REL 28, LAST SEQUENCE UPDATE)
DI 15-FEB-1999 (REL 35, LAST ANNOTATION UPDATE)
DE RAL-LACTAMASE 1 PRECURSOR (EC 3.2.2) (PENICILLINASE).
GN BLAL.
OS STREPTOMYCES CAZALI.
OC BACTERIA: FIRMICUTES: ACTINOBACTERIA: ACTINOMYCETIDAE:
OC ACTINOMYCETALES: STREPTOMYCETACEAE: STREPTOMYCETIDAE:
Query Match 100.0% Score 39 DB 1 Length 292
Rest Local Similarity 100.0% Prid No 5 466-00
Matches 5 Conservative 0 Mismatches 0 Indels 0 Gaps 0
Db 161 VWRPP 165
11111
QY 1 VWRPP 5
RESULT 4
ID YHT-ECOLI STANDARD: PRT: 292 AA.
AC P32141:
DI 01-OCT-1993 (REL 27, CREATED)
DI 01-OCT-1993 (REL 27, LAST SEQUENCE UPDATE)
DI 01-NOV-1997 (REL 35, LAST ANNOTATION UPDATE)
DE YHT-ECOLI 22 KD PROTEIN IN GLNA-SEN INTERGENIC REGION (F033).
GN YHT.
OS ESCHERICHIA COLI.
OC BACTERIA: PROTEOBACTERIA: GAMMA SUBDIVISION: ENTEROBACTERIACEAE:
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655:
RX MEDLINE: 9347969.
PA PIONNETT G., HILLIARD V.D., DANFELS D.L., PIATNER P.:
RI "Analysis of the Escherichia coli genome. II: DNA sequence of the
RI region from 87.2 to 89.2 minutes."
RI NUCLEIC ACIDS RES. 21:3391-3399(1993).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.ist-sib.ch/announcements
CC or send an email to license@sib-sib.ch.)
CC -----
DE EMBL: I32201; GI: 330495;
DE EMBL: AB000464; GI: 740314;
DE PIR: S40925; S40925.
DE PIR: F011846; YHT
KW HYPOTHETICAL PROTEIN.
SI SEQUENCE 292 AA: 31983 MW: 60818471 QRC32:
Query Match 100.0% Score 39 DB 1 Length 292
Rest Local Similarity 100.0% Prid No 5 466-00
Matches 5 Conservative 0 Mismatches 0 Indels 0 Gaps 0
Db 161 VWRPP 165
11111
QY 1 VWRPP 5

```



```

KW      3D-STRUCTURE.      1      15
FT SIGNAL                16      549
FT CHAIN                 224     356 BY SIMILARITY. LIPASE 3.
FT ACT_SITE              464     464 BY SIMILARITY.
FT ACT_SITE              464     464 BY SIMILARITY.
FT ACT_SITE              464     464 BY SIMILARITY.
FT DISULFIDE             280     320 POTENTIAL.
FT DISULFIDE             280     320 POTENTIAL.
FT CAPOXYR               466     555
FT CANOXYR               466     555
SQ SEQUENCE 549 AA; 52754 MW; 5840800 DPT320;

Query Match          100 %; Score 347.16 (E 1); Length 549;
Seq-Local Similarity 100 %; Field No. 5.46e+00;
Matches           5; Conservative    0; Mismatches   0; Gaps       0; Gaps

Db      117 VWRPP 121
QY      1 VWRPP 5

RESULT          9
AC ID LIP_CANRO STANDARD; PRT: 549 AA.
AD F00261;
BT 01-FEB-1991 (PRL 17, CREATED)
DT 01-MAY-1993 (PRL 27, LAST SEQUENCE UPDATE)
DE 15-FEB-1996 (PRL 37, LAST ANNOTATION UPDATE)
DE LIPASE 1 PRECURSOR (EC 3.1.1.3).
GN LIPI.
OS CANDIDA RUJOSA (YEAST) [CANDIDA TROPICALEA];
OC EUKARYOTA; FUNGI; ASCOMYCOTA; REMIASCOMICETES; SACCHAROMYCETALES;
OC CANTILLACEAE; CANDIDA.
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14830;
RX MEDLINE: 92305068.
RA LONGHI S., FUSELLI F., BRANDINI R., LOTTI M., VACCINI M.,
RA AIREPUGHINA L.;
PT "Cloning and nucleotide sequence of the lipase gene from Candida cylindracea";
PF RITCHIEIM RICHIPS ACTA 1193 227 232(1992).
RN [2]
RF SEQUENCE OF 12,542 FRAM N.A. AMT PARTIAL SEQUENCE.
RC STRAIN=ATCC 14830 / MS-5;
RX MEDLINE. 89384874.
RA KAWAGUCHI Y., HONDA H., TANIGUCHI-MORIMURA Y., IWASAKI S.;
PT "The cyp36B gene is read as surface in an aspartic acid yeast Candida cylindracea";
PF RITCHIEIM RICHIPS ACTA 1193 227 232(1992).
RN [3]
RF X-RAY CRYSTALLOGRAPHY (2.06 ANGSTROMS).
PX MEDLINE: 92986131.
RA GRUBBSHUIJ F., LI Y., SCHWAB J.D., BOUTHERLIER F., SMITH P.,
RA HARRISON D., RUBIN E., OGLER M.J.;
PT "Insights into haemoglobin activation from x-ray crystallographic studies of Candida rugosa lipase";
PF RITCHIEIM RICHIPS ACTA 1193 227 232(1992).
RN [4]
RF X-RAY CRYSTALLOGRAPHY (3.05 ANGSTROMS).
PX MEDLINE: 94190867.
RA GRUBBSHUIJ F., BOUTHERLIER F., VANLAEREKAS F., REBERT A.M.,
FA STRAS J.P., PRIMER E., BUTLER M.;
PT "Analogs of reaction intermediates identify a unique substrate binding site in Candida rugosa lipase.";
PF RITCHIEIM RICHIPS ACTA 1193 227 232(1992).
RN [5]
RP REVIEW
RX MEDLINE: 98451816.
RA BENJAMIN S., PANDEY A.;
PT "Candida rugosa lipase: molecular biology and versatility in biotechnology.";
RN YEAST 14:1069-1087(1998).
RA YEAST 14:1069-1087(1998).

```

11 A FAMILY AND AN IN...  
 12 ...  
 13 ...  
 14 ...  
 15 ...  
 16 ...  
 17 ...  
 18 ...  
 19 ...  
 20 ...  
 21 ...  
 22 ...  
 23 ...  
 24 ...  
 25 ...  
 26 ...  
 27 ...  
 28 ...  
 29 ...  
 30 ...  
 31 ...  
 32 ...  
 33 ...  
 34 ...  
 35 ...  
 36 ...  
 37 ...  
 38 ...  
 39 ...  
 40 ...  
 41 ...  
 42 ...  
 43 ...  
 44 ...  
 45 ...  
 46 ...  
 47 ...  
 48 ...  
 49 ...  
 50 ...  
 51 ...  
 52 ...  
 53 ...  
 54 ...  
 55 ...  
 56 ...  
 57 ...  
 58 ...  
 59 ...  
 60 ...  
 61 ...  
 62 ...  
 63 ...  
 64 ...  
 65 ...  
 66 ...  
 67 ...  
 68 ...  
 69 ...  
 70 ...  
 71 ...  
 72 ...  
 73 ...  
 74 ...  
 75 ...  
 76 ...  
 77 ...  
 78 ...  
 79 ...  
 80 ...  
 81 ...  
 82 ...  
 83 ...  
 84 ...  
 85 ...  
 86 ...  
 87 ...  
 88 ...  
 89 ...  
 90 ...  
 91 ...  
 92 ...  
 93 ...  
 94 ...  
 95 ...  
 96 ...  
 97 ...  
 98 ...  
 99 ...  
 100 ...

11 TURN 224 224  
 12 HELIX 225 224  
 13 STRAND 242 243  
 14 TURN 244 245  
 15 STRAND 246 247  
 16 STRAND 251 255  
 17 TURN 255 256  
 18 HELIX 256 281  
 19 TURN 282 282  
 20 STRAND 284 285  
 21 TURN 289 295  
 22 HELIX 298 305  
 23 TURN 306 307  
 24 TURN 311 312  
 25 TURN 314 317  
 26 HELIX 334 334  
 27 TURN 341 341  
 28 STRAND 348 353  
 29 TURN 354 354  
 30 STRAND 355 355  
 31 TURN 356 357  
 32 HELIX 358 364  
 33 TURN 365 365  
 34 HELIX 370 380  
 35 TURN 382 383  
 36 HELIX 386 395  
 37 TURN 400 401  
 38 TURN 406 407  
 39 TURN 409 412  
 40 TURN 416 417  
 41 HELIX 418 429  
 42 TURN 430 430  
 43 HELIX 431 440  
 44 STRAND 446 451  
 45 TURN 453 456  
 46 TURN 458 460  
 47 STRAND 463 463  
 48 TURN 464 465  
 49 HELIX 466 472  
 50 TURN 473 473  
 51 TURN 477 478  
 52 HELIX 479 482  
 53 TURN 483 483  
 54 HELIX 484 492  
 55 TURN 495 496  
 56 STRAND 507 507  
 57 TURN 510 511  
 58 STRAND 517 520  
 59 STRAND 525 528  
 60 TURN 532 533  
 61 HELIX 534 541  
 62 HELIX 544 547  
 63 SEQUENCE 549 AA: 5955 MW: 67945.69 Da

Query Match 100.0% Score 49; Da 1; Length 549  
 Best Local Similarity 100.0% Pred. No. 5.46e-004  
 Matches 7; Conserved 1; Mismatches 1; Gaps 1

DB 117 VVRRP 121  
 Q7 1 VVRRP 5

RESULT 10  
 ID LIP5\_CANP0 STANFORD PPT: 549 AA  
 AC P32493  
 DT 01-OCT-1994 (REL. 27, CREATED)  
 DT 01-OCT-1994 (REL. 27, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 47, LAST ANNOTATION UPDATE)  
 DE LIPASE 5 PRECURSOR (EC 3.1.1.7).  
 GN LIP5  
 OS CANDIDA K1070SA (YEAST) (CANDIDA CYLINDRACA)  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES



```

13
14 YLUN TAPEL STANDARD FRT 748 AA.
15
16 15-JUL-1998 (REL. 46, CHEATER)
17 15-JUL-1998 (REL. 46, LAST SEQUENCE UPDATE)
18 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
19 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
20 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
21 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
22 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
23 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
24 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
25 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
26 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
27 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
28 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
29 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
30 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
31 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
32 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
33 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
34 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
35 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
36 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
37 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
38 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
39 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
40 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
41 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
42 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
43 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
44 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
45 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
46 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
47 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
48 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
49 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
50 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
51 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
52 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
53 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
54 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
55 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
56 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
57 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
58 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
59 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
60 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
61 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
62 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
63 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
64 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
65 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
66 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
67 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
68 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
69 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
70 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
71 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
72 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
73 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
74 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
75 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
76 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
77 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
78 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
79 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
80 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
81 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
82 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
83 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
84 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
85 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
86 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
87 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
88 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
89 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
90 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
91 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
92 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
93 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
94 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
95 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
96 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
97 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
98 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
99 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)
100 15-JUL-1998 (REL. 46, LAST ANNOTATION UPDATE)

```

```

101 679 VVRPP 684
102 1 VVRPP 5
103 1 VVRPP 5
104 1 VVRPP 5
105 1 VVRPP 5
106 1 VVRPP 5
107 1 VVRPP 5
108 1 VVRPP 5
109 1 VVRPP 5
110 1 VVRPP 5
111 1 VVRPP 5
112 1 VVRPP 5
113 1 VVRPP 5
114 1 VVRPP 5
115 1 VVRPP 5
116 1 VVRPP 5
117 1 VVRPP 5
118 1 VVRPP 5
119 1 VVRPP 5
120 1 VVRPP 5
121 1 VVRPP 5
122 1 VVRPP 5
123 1 VVRPP 5
124 1 VVRPP 5
125 1 VVRPP 5
126 1 VVRPP 5
127 1 VVRPP 5
128 1 VVRPP 5
129 1 VVRPP 5
130 1 VVRPP 5
131 1 VVRPP 5
132 1 VVRPP 5
133 1 VVRPP 5
134 1 VVRPP 5
135 1 VVRPP 5
136 1 VVRPP 5
137 1 VVRPP 5
138 1 VVRPP 5
139 1 VVRPP 5
140 1 VVRPP 5
141 1 VVRPP 5
142 1 VVRPP 5
143 1 VVRPP 5
144 1 VVRPP 5
145 1 VVRPP 5
146 1 VVRPP 5
147 1 VVRPP 5
148 1 VVRPP 5
149 1 VVRPP 5
150 1 VVRPP 5
151 1 VVRPP 5
152 1 VVRPP 5
153 1 VVRPP 5
154 1 VVRPP 5
155 1 VVRPP 5
156 1 VVRPP 5
157 1 VVRPP 5
158 1 VVRPP 5
159 1 VVRPP 5
160 1 VVRPP 5
161 1 VVRPP 5
162 1 VVRPP 5
163 1 VVRPP 5
164 1 VVRPP 5
165 1 VVRPP 5
166 1 VVRPP 5
167 1 VVRPP 5
168 1 VVRPP 5
169 1 VVRPP 5
170 1 VVRPP 5
171 1 VVRPP 5
172 1 VVRPP 5
173 1 VVRPP 5
174 1 VVRPP 5
175 1 VVRPP 5
176 1 VVRPP 5
177 1 VVRPP 5
178 1 VVRPP 5
179 1 VVRPP 5
180 1 VVRPP 5
181 1 VVRPP 5
182 1 VVRPP 5
183 1 VVRPP 5
184 1 VVRPP 5
185 1 VVRPP 5
186 1 VVRPP 5
187 1 VVRPP 5
188 1 VVRPP 5
189 1 VVRPP 5
190 1 VVRPP 5
191 1 VVRPP 5
192 1 VVRPP 5
193 1 VVRPP 5
194 1 VVRPP 5
195 1 VVRPP 5
196 1 VVRPP 5
197 1 VVRPP 5
198 1 VVRPP 5
199 1 VVRPP 5
200 1 VVRPP 5

```

This SWISS PRO entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

This SWISS PRO entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

Fri Feb 25 14:06:58 2000

US-09-132-799-2.rsp

Page 8

Search completed: Fri Feb 25 13:17:15 2000  
Job time : 47 secs.



\*\*\*\*\*  
WATERMAN  
(TM)  
\*\*\*\*\*  
Release 3.1A John F. Collins, Biocomputing Research Unit,  
Copyright (c) 1994-1998 University of Edinburgh, G. K.  
Distribution rights by Oxford Molecular Ltd

MPsr-dupff protein - protein database search, using Smith-Waterman algorithm  
Run on: Fri Feb 25 13:17:33 2000: MasPar time 4.44 Seconds  
Laboratory output not generated. 61,508 Million cell updates/sec

1. Job: 095-132-799-2  
Description: (1.5) from US-09-132-799.psp  
Database: 43  
Sequence: 1 VVSEF 5  
Sequence: 1000  
Seq. desc: 1000 seqs, 54570741 residues  
Post processing: MasPar March 1999  
Lasted first 1000 summaries  
1. Job: 095-132-799-2  
Description: (1.5) from US-09-132-799.psp  
Database: 43  
Sequence: 1 VVSEF 5  
Sequence: 1000  
Seq. desc: 1000 seqs, 54570741 residues  
Post processing: MasPar March 1999  
Lasted first 1000 summaries

Statistics: Mean 16.974; Variance 16.567; scale 1.139

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
It is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description	Prod. No.
1	49	100.0	63	6	UNKN-WN PROTEIN (FRAGM)	8.83e+00
2	49	100.0	82	2	CRCA-1, CRGB.	8.83e+00
3	49	100.0	115	10	CROCEFERIN TYPE C (FRA	8.83e+00
4	49	100.0	126	5	F2155.8 PROTEIN.	8.83e+00
5	49	100.0	145	2	MAV145.	8.83e+00
6	49	100.0	163	10	CROCEFERIN (12S STORAG	8.83e+00
7	49	100.0	166	10	CROCEFERIN (FRAGMENT).	8.83e+00
8	49	100.0	200	14	COAT PROTEIN.	8.83e+00
9	49	100.0	212	14	COAT PROTEIN.	8.83e+00
10	49	100.0	213	10	PROLINE RICH PROTEIN.	8.83e+00
11	49	100.0	219	10	MTM4 (FRAGMENT).	8.83e+00
12	49	100.0	274	10	PUTATIVE MEMBRANE GLYC	8.83e+00
13	49	100.0	324	2	KBPK.	8.83e+00
14	49	100.0	327	2	AMINOMETHYLTRANSFERASE	8.83e+00
15	49	100.0	371	10	GIRRELLIN C-20 OXIDA	8.83e+00
16	49	100.0	407	10	CELL WALL PROTEIN.	8.83e+00
17	49	100.0	407	5	HYDROLYTIC ENZYME.	8.83e+00
18	49	100.0	409	14	HYPOTHETICAL 45.5 KD P	8.83e+00
19	49	100.0	421	5	SIMILAR TO POTASSIUM C	8.83e+00
20	49	100.0	421	5	SIMILAR TO POTASSIUM C	8.83e+00

21	39	100.0	444	2	CHEM TAXIS METER TEE	8.83e+00
22	39	100.0	449	10	127A16.30 PROTEIN (FRA	8.83e+00
23	39	100.0	476	10	GLUCONIN PEPTIDES 4	8.83e+00
24	39	100.0	477	10	INT 10-1 A UNILATE BUA	8.83e+00
25	39	100.0	481	10	PG12.4 PROTEIN.	8.83e+00
26	39	100.0	481	10	PG12.4 PROTEIN.	8.83e+00
27	39	100.0	506	10	CROCEFERIN (12S STORAG	8.83e+00
28	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
29	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
30	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
31	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
32	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
33	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
34	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
35	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
36	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
37	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
38	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
39	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
40	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
41	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
42	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
43	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
44	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
45	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
46	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
47	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
48	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
49	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
50	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
51	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
52	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
53	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
54	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
55	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
56	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
57	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
58	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
59	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
60	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
61	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00
62	39	100.0	506	10	CROCEFERIN (FRAGMENT).	8.83e+00

Note: Post processor removed 149 summaries from list due to seed 0. Parameters used:

ALIGNMENTS

RESULT	1	PRELIMINARY	1P1	1A
ID	Q29255			1A
AC	Q29255			1A
DT	01-NOV-1996 (TREMELIN, 01. CREATED)			
DI	01-NOV-1996 (TREMELIN, 01. LAST SEQUENCE UPDATE)			
DI	01-MAY-1997 (TREMELIN, 01. LAST SEQUENCE UPDATE)			
DE	UNKN-WN PROTEIN (FRAGMENT).			
OS	SUS SCROGA (F16).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA			
OC	ARTIODACTYLA; SUINIFORMES; SUINA; SUIIDAE; SUI.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-SMALL INTESTINE			
RA	WINFRED A.K. FREDERICK M. DAVIES W.			
RL	EMBL: GENE 71504-517 (1996)			
DR	EMBL: F15071: E196745			
FT	NON-TER 1			
FT	NON-TER 63			
SQ	SEQUENCE 63 AA: 4443 MK: 80RZA707 GRC92			

Query Match: 100.0% Score 39. ID 61. Length 63  
Best Local Similarity 100.0% E-Value 0.000000  
Database: E. coli, Mus musculus, R. norvegicus, S. cerevisiae

DR 51 VVR09 55

|||||  
QY 1 VVRPP 5

RESULT 2  
ID 007848  
AC 007848  
DT 01-NOV-1997 (TEMPREL 04, CREATED)  
DT 01-JUL-1997 (TEMPREL 04, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (TEMPREL 04, LAST SEQUENCE UPDATE)  
DE CPCC  
GN CROB  
OS ESCHERICHIA COLI  
OC BACTERIA: PROTEOBACTERIA: GAMMA SUBDIVISION: ENTEROBACTERIAEAL  
OC ESCHERICHIA  
RN (1)  
RP SEQUENCE FROM N.A.  
PV WFLINE 0700112  
RA HU Y H, HU F, DEAN R, GINERAS M, PETRAFF W, TEUN N J  
FT "Genetic analysis of three genes located in the resistance and  
RI Chromosome localization in Escherichia coli"  
RL GENETICS 143:1521-1532(1996)  
DR EMBL: S83896; E323898  
SQ SEQUENCE 82 AA, 2024 MW, P295427, Q6012

Query Match 100.0% Score 24, 28.5, Length 115  
Best Local Similarity 100.0% Pred No. 8, 84e+00;  
Matches 5: Conservative 0, Mismatches 0, Indels 0, Gaps 0

Db 35 VVRPP 39  
QY 1 VVRPP 5

RESULT 3  
ID 042181  
AC 042181  
DT 01-NOV-1996 (TEMPREL 01, CREATED)  
DT 01-NOV-1996 (TEMPREL 01, LAST SEQUENCE UPDATE)  
DT 01-AUG-1998 (TEMPREL 07, LAST ANNOTATION UPDATE)  
DE CROCIFERIN TYPE C (FRAGMENT)  
GN ARABIDOPSIS THALIANA (MOUSE-EAR CRESS)  
OC EUKARYOTA: VIRIDIPHYTES: EUPHYCOPHYTES: EUPHYCOPHYTES: TRACHEOPHYTES  
OC EUPHYCOPHYTES: SPERMATOPHYTES: MAGNOLIOPHYTES: EUPHYCOPHYTES: ROSALES  
OC CAPARALES: BRASSICACEAE: ARABIDOPSIS  
RN (1)  
RP SEQUENCE FROM N.A.  
RC TISSUE-DRY SEEDS OF A THALIANA ECOTYPE COLUMBIA:  
RA RAYAL M, GELLET F, LACIE M, MEYER T, SCORR R, DELSENY M,  
KL SUBMITTED (NOV-1993) TO EMBL/GENBANK/CCRB DATA BANK  
DR EMBL: 0417512  
FT NON-TER 1  
FT NON-TER 1  
SQ SEQUENCE 115 AA, 13019 MW, R10666, Q6032

Query Match 100.0% Score 32, 38.12, Length 117  
Best Local Similarity 100.0% Pred No. 8, 84e+00;  
Matches 5: Conservative 0, Mismatches 0, Indels 0, Gaps 0

Db 84 VVRPP 88  
QY 1 VVRPP 5

RESULT 4  
ID 019685  
AC 019685  
DT 01-NOV-1996 (TEMPREL 01, CREATED)  
DT 01-NOV-1996 (TEMPREL 01, LAST SEQUENCE UPDATE)  
DT 01-JAN-1997 (TEMPREL 03, LAST ANNOTATION UPDATE)  
DE F2105 8 PROTEIN  
GN CAENORHABDITIS ELEGANS  
OC

EUKARYOTA: METAZOA: NEMATODA: SECERNENTEA: RHABDITIA: RHABDITIDA:  
OC RHABDITIA: RHABDITIDAE: RHABDITIDAE: RHABDITIDAE: CAENORHABDITIS  
RN (1)  
RP SEQUENCE FROM N.A.  
RA BERKS M J  
RC "Sequencing of the 18S rRNA gene from the nematode, Caenorhabditis  
RN (2)  
RP SEQUENCE FROM N.A.  
RA MEDLINE: 94150718  
RA WILSON P, AINSWORTH P, ANTERSON K, RAYNES C, BERKS M,  
RA ROSEFIELD J, BORTON J, CONNELL M, CUPSEY I, COOPER T, GUNSON A,  
RA CRAXIN M, DEAR S, DE J, EUBIN E, FAVELLO A, FULTON L,  
RA GARNER A, GREEN P, HARKINS L, HILLIER N, JEFF M, JOHNSTON L,  
RA JONES M, KERSHAN J, KIRSTEN J, LAISTER N, PATRILL P,  
RA LIGHTNING J, LLOYD C, MCMURRAY A, MORTIMORE R, O'CALLAGHAN M,  
RA PARSONS J, PERCY C, RIEKEN L, ROOPRA A, SAUNDERS D, SHOWNKEEN P,  
RA SMAILDON N, SMITH A, SONNHAMMER E, STADEN P, SLESTON J,  
RA THERRY-MIEG T, THOMAS P, VAUGHAN K, WATERSTON P,  
RA WATSON A, WELSHOP L, WILKINSON-SPECKAT J, WHITMAN P,  
FT "2.2 kb of the 18S rRNA gene from the nematode, Caenorhabditis  
PL NATURE 368:32-39(1994)  
DR EMBL: 254271; E1348020  
SQ SEQUENCE 105 AA, 34502 MW, P56000, Q6022

Query Match 100.0% Score 32, 38.5, Length 126  
Best Local Similarity 100.0% Pred No. 8, 84e+00;  
Matches 5: Conservative 0, Mismatches 0, Indels 0, Gaps 0

Db 109 VVRPP 113  
QY 1 VVRPP 5

RESULT 5  
ID 007404  
AC 007404  
DT 01-NOV-1997 (TEMPREL 04, CREATED)  
DT 01-JUL-1997 (TEMPREL 04, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TEMPREL 08, LAST ANNOTATION UPDATE)  
DE MAV145  
GN MYCOBACTERIUM AVIUM  
OC MYCOBACTERIUM AVIUM  
OC BACTERIA: FIRMICUTES: ACTINOBACTERIA: ACTINOBACTERIAE  
OC ACTINOMYCETIALES: CORYNEBACTERIACEAE: MYCOBACTERIACEAE: MYCOBACTERIUM  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STAIN-GIP10:  
RA MEDLINE: 9335739  
RA LAM M, GUSSEPH I, DE POST F, SERRAIF P, RICHART G,  
FT "Identification of a 145 bp variable sequence around the inhA gene  
FT of Mycobacterium avium and similarity analysis of the products of  
FT putative ORFs"  
RL MICROBIOLOGY 144:807-814(1998)  
DR EMBL: AF002233; G182223  
SQ SEQUENCE 145 AA, 31550 MW, P56000, Q6022

Query Match 100.0% Score 32, 38.5, Length 145  
Best Local Similarity 100.0% Pred No. 8, 84e+00;  
Matches 5: Conservative 0, Mismatches 0, Indels 0, Gaps 0

Db 64 VVRPP 68  
QY 1 VVRPP 5

RESULT 6  
ID 002503  
AC 002503  
DT 01-NOV-1996 (TEMPREL 01, CREATED)  
DT 01-NOV-1996 (TEMPREL 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TEMPREL 08, LAST ANNOTATION UPDATE)  
DE CROCIFERIN (12S STORAGE PROTEIN) (FRAGMENT)



DE CRUCIFERIN (FRAGMENT).  
 GN PAP7.  
 OS PAPHANUS SATIVUS (RADISH).  
 CC EUPHYLLOPHYTES: VIRIDIPALANAE: SIREPTOPHYTES: EMBRYOPHYTES: TRACHEOPHYTES:  
 CC EUPHYLLOPHYTES: SPERMATOPHYTES: MAGNOLIOPHYTES: EUDICOTYLEONS: ROSIDAE:  
 CC CAPPARALES: BRASSICACEAE: RAPANUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=CV JEMALONG 35; TISSUE=ROOT NODULE;  
 RA GAMES P.1  
 RA STEINER (1997), 1. FUEL, RESEARCH, 1997 DATA BANKS.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=CV JEMALONG 35; TISSUE=ROOT NODULE;  
 RA HELMEL, 2002024.  
 RA GAMES P.1. IF CARVALHO NITEL F. (1998) N. (PHILIPPO J.).  
 RA Use of a subtractive hybridization approach to identify new Malvaceae  
 RA truncatula genes induced during root nodule development.  
 RA MOL. PLANT-MICROBE INTERACT. 9:223-242(1996).  
 DR EMBL: V15372; E1170495; .  
 FI NON-ITER 1  
 FI SEQUENCE 242 AA, 25223 MW, 55115523 25232;  
 SQ SEQUENCE 242 AA, 25223 MW, 55115523 25232;  
 Query Match 100.0%; Score 39; DB 10; Length 233;  
 Best Local Similarity 100.0%; Pred. No. 8.83e-00;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 16 VVRPP 20  
 II  
 QY 1 VVRPP 5  
 RESULT 11  
 ID Q39686 PRELIMINARY PRT 245 AA  
 AC Q39686  
 DT 01-NOV-1996 (TEMRELPEL 01, CREATED)  
 DI 01-NOV-1996 (TEMRELPEL 01, LAST SEQUENCE UPDATE)  
 DI 01-JAN-1998 (TEMRELPEL 02, LAST SEQUENCE UPDATE)  
 DE PROLINE-RICH PROTEIN  
 CC DACTYLIS CAROTA (CARROT).  
 CC EUPHYLLOPHYTES: VIRIDIPALANAE: SIREPTOPHYTES: EMBRYOPHYTES: TRACHEOPHYTES:  
 CC EUPHYLLOPHYTES: SPERMATOPHYTES: MAGNOLIOPHYTES: EUDICOTYLEONS:  
 CC ASTERIDAE: ARALIAE: APIACEAE: DAUCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MERLIN, 2410293  
 RA FRUTER W. FOWLER I. J., SUZUKI H., SHAVER G., TIERNEY M. L.  
 RA "Expression of DCP1 is linked to carbon storage, root formation and  
 RA root growth by a 3.1 kb 3'5' regulatory region".  
 RL PLANT PHYSIOL. 101:259-266(1993).  
 DR EMBL: X59802; G01106; .  
 FI NON-ITER 1  
 FI SEQUENCE 245 AA, 25223 MW, 55062607 25232;  
 SQ SEQUENCE 245 AA, 25223 MW, 55062607 25232;  
 Query Match 100.0%; Score 39; DB 10; Length 235;  
 Best Local Similarity 100.0%; Pred. No. 8.83e-00;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

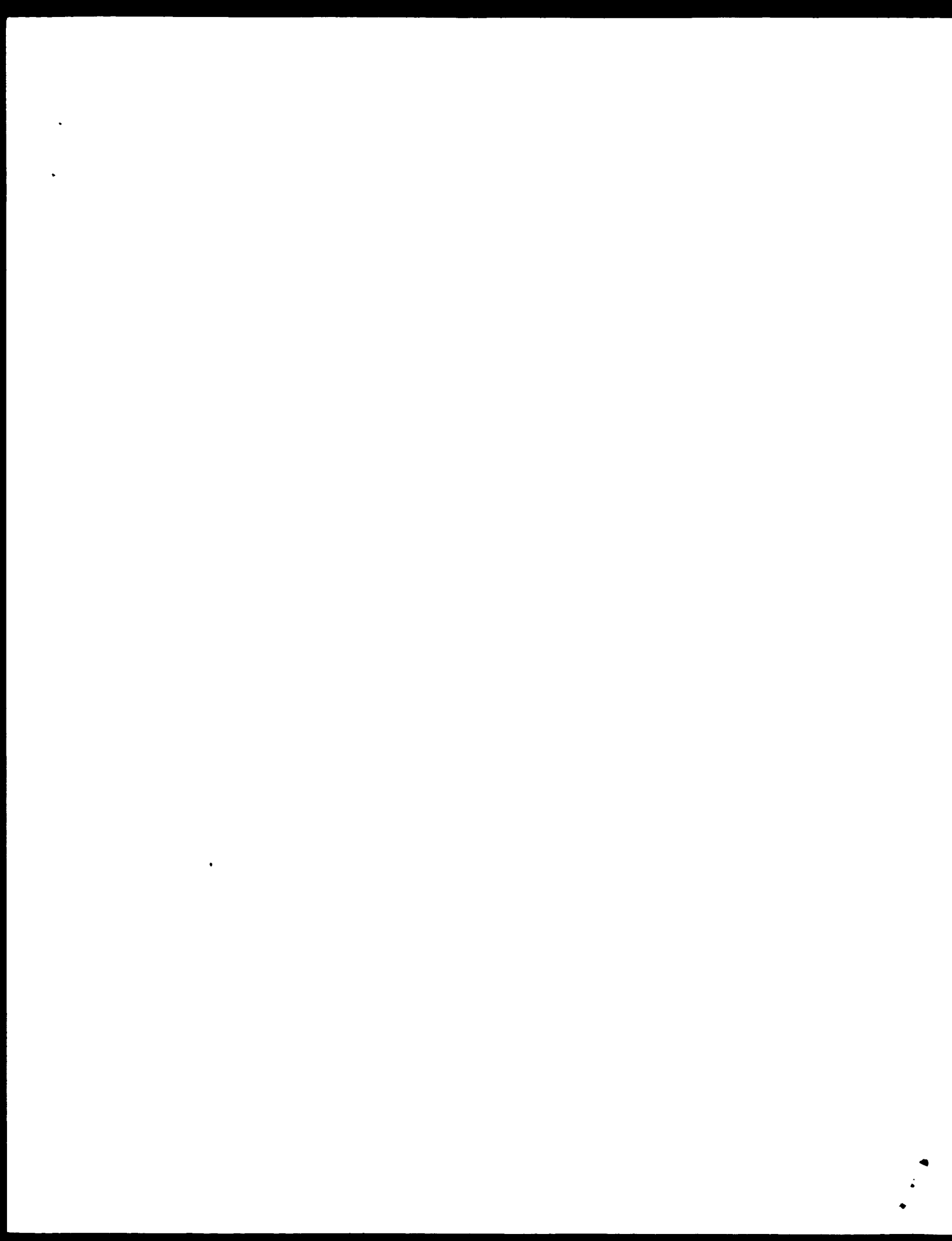
DB 208 VVRPP 212  
 II  
 QY 1 VVRPP 5  
 RESULT 12  
 ID Q24102 PRELIMINARY PRT 245 AA  
 AC Q24102  
 DT 01-JAN-1998 (TEMRELPEL 05, CREATED)  
 DI 01-JAN-1998 (TEMRELPEL 05, LAST SEQUENCE UPDATE)  
 DI 01-NOV-1996 (TEMRELPEL 02, LAST SEQUENCE UPDATE)  
 DE MTN4 (FRAGMENT).  
 RN [1]  
 RN MTN4

OS MEDICAGO TRUNCATULA (BARREL MEDIC).  
 CC EUKARYOTA: VIRIDIPALANAE: SIREPTOPHYTES: EMBRYOPHYTES: TRACHEOPHYTES:  
 CC EUPHYLLOPHYTES: SPERMATOPHYTES: MAGNOLIOPHYTES: EUDICOTYLEONS: ROSIDAE:  
 CC FABALES: FABACEAE: PAPILIONOIDEAE: MEDICAGO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=CV JEMALONG 35; TISSUE=ROOT NODULE;  
 RA GAMES P.1  
 RA STEINER (1997), 1. FUEL, RESEARCH, 1997 DATA BANKS.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=CV JEMALONG 35; TISSUE=ROOT NODULE;  
 RA HELMEL, 2002024.  
 RA GAMES P.1. IF CARVALHO NITEL F. (1998) N. (PHILIPPO J.).  
 RA Use of a subtractive hybridization approach to identify new Malvaceae  
 RA truncatula genes induced during root nodule development.  
 RA MOL. PLANT-MICROBE INTERACT. 9:223-242(1996).  
 DR EMBL: V15372; E1170495; .  
 FI NON-ITER 1  
 FI SEQUENCE 242 AA, 25223 MW, 55115523 25232;  
 SQ SEQUENCE 242 AA, 25223 MW, 55115523 25232;  
 Query Match 100.0%; Score 39; DB 10; Length 240;  
 Best Local Similarity 100.0%; Pred. No. 8.83e-00;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 51 VVRPP 55  
 II  
 QY 1 VVRPP 5  
 RESULT 13  
 ID Q82371 PRELIMINARY PRT 274 AA  
 AC Q82371  
 DT 01-NOV-1996 (TEMRELPEL 02, CREATED)  
 DI 01-NOV-1996 (TEMRELPEL 02, LAST SEQUENCE UPDATE)  
 DI 01-NOV-1996 (TEMRELPEL 02, LAST SEQUENCE UPDATE)  
 DE POLYMERASE CHAIN REACTION  
 CC DACTYLIS CAROTA (CARROT).  
 CC EUPHYLLOPHYTES: VIRIDIPALANAE: SIREPTOPHYTES: EMBRYOPHYTES: TRACHEOPHYTES:  
 CC EUPHYLLOPHYTES: SPERMATOPHYTES: MAGNOLIOPHYTES: EUDICOTYLEONS: ROSIDAE:  
 CC FABALES: BRASSICACEAE: ARABIDOPSIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=CV COLUMBIA  
 RA ROSENLEY S. B. LIN X., KAUL S., SHEA T. P., FUDIT G. Y., MASON T. M.,  
 RA SHEN M., RICHARDSON M., FRASER J. M., SUMERVILLE J. P., VENTER T. C.,  
 RA "Arabidopsis thaliana chromosome 11 RAC 127A16 genomic sequence".  
 RA SUBMITTED (SEP-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL: A0005466; G1602319; .  
 FI NON-ITER 1  
 FI SEQUENCE 274 AA, 28791 MW, 55106111 28791;  
 SQ SEQUENCE 274 AA, 28791 MW, 55106111 28791;  
 Query Match 100.0%; Score 39; DB 10; Length 274;  
 Best Local Similarity 100.0%; Pred. No. 8.83e-00;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 175 VVRPP 179  
 II  
 QY 1 VVRPP 5  
 RESULT 14  
 ID P72144 PRELIMINARY PRT 320 AA  
 AC P72144  
 DT 01-FEB-1997 (TEMRELPEL 02, CREATED)  
 DI 01-FEB-1997 (TEMRELPEL 02, LAST SEQUENCE UPDATE)  
 DI 01-NOV-1996 (TEMRELPEL 02, LAST SEQUENCE UPDATE)  
 DE WBBK.  
 RN WBBK.  
 OS PSEUDOMONAS AERUGINOSA.  
 CC BACTERIA: FRIGIDIBACTERIA: GAMMA SUBDIVISION: PSEUDOMONAS GROUP:  
 CC PSEUDOMONAS.  
 RN [1]







Fri Feb 25 14:07:03 2000

US-09-132-799-3.rag

Page 2

Search completed: Fri Feb 25 14:07:03 2000  
Job time : 80 secs.

---







\*\*\*\*\*

release 3.1A John F. Collins, Bioinformatics Research Unit,  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

\*\*\*\*\*

MSearch: protein - protein database search, using Smith-Waterman algorithm

Parameters: File: P3 25 13.24.07 2000. MaxPar time 3.05 seconds 78.813 Million cell updates/sec

Database: P3 25 13.24.07 2000. MaxPar time 3.05 seconds 78.813 Million cell updates/sec

Description: 4036854 residues  
 Database: P3 25 13.24.07 2000. MaxPar time 3.05 seconds 78.813 Million cell updates/sec

Database: P3 25 13.24.07 2000. MaxPar time 3.05 seconds 78.813 Million cell updates/sec

Database: P3 25 13.24.07 2000. MaxPar time 3.05 seconds 78.813 Million cell updates/sec

Database: P3 25 13.24.07 2000. MaxPar time 3.05 seconds 78.813 Million cell updates/sec

Database: P3 25 13.24.07 2000. MaxPar time 3.05 seconds 78.813 Million cell updates/sec

Database: P3 25 13.24.07 2000. MaxPar time 3.05 seconds 78.813 Million cell updates/sec

Prod. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is defined by analysis of the total score distribution.

# SUMMARIES

Result	Score	Query	Match	Length	DB ID	Description	Prod. No.
1	47	1-172	172	2	S28082	vif protein - simian	2.01e+00
2	47	1-172	211	2	R23162	3xK proline-rich prot	2.01e+00
3	47	1-172	236	2	J02260	proline rich cell wal	2.01e+00
4	46	1-172	99	2	S28083	cell wall proline-rich	5.48e+00
5	46	1-172	449	2	J14081	glyceroldehyde 3 phos	5.48e+00
6	46	1-172	449	2	J07147	probable histone H4-ace	5.48e+00

N.B. P3 25 13.24.07 2000. MaxPar time 3.05 seconds 78.813 Million cell updates/sec

## ALIGNMENTS

Result 1  
 Entry: S28082  
 Title: vif protein - simian immunodeficiency virus  
 Organism: simian immunodeficiency virus, SIV  
 Date: 17-Apr-1993  
 Accession: J02260  
 Molecule type: DNA  
 Residues: 1-172  
 Cross-references: P01477; N01477; N01477-24  
 Experimental source: vat. Nates  
 Comment: This protein is a wall and is involved in cell  
 Classification: #superfamily proline-rich protein  
 Keywords: cell wall

Journal: Nature (1989) 341:549-551  
 Title: Wild-catch? African mandrill  
 Cross-references: M010601676  
 Accession: S28082  
 Status: nucleic acid sequence not shown; translation not shown  
 Molecule type: DNA  
 Residues: 1-172  
 Cross-references: P01477; N01477; N01477-24  
 Comment: This protein is a wall and is involved in cell  
 Classification: #superfamily proline-rich protein  
 Keywords: cell wall

## GENETICS

Gene: vif  
 Search: 172  
 Query Match: 100.0%  
 Best Local Similarity: 100.0%  
 Matches: 6; Conservative: 6; Mismatches: 6; Indels: 6; Gaps: 6

DB: 95 VWRP 100  
 QY: 1 VWRP 6

## RESULT

Entry: R23162  
 Title: 3xK proline-rich protein - carrot (fragment)  
 Organism: Formal name: Nicotiana glauca; common name: carrot  
 Date: 06-Aug-1987  
 Accession: J02260  
 Molecule type: mRNA  
 Residues: 1-211  
 Cross-references: P01477; N01477; N01477-24  
 Experimental source: vat. Nates  
 Comment: This protein is a wall and is involved in cell  
 Classification: #superfamily proline-rich protein  
 Keywords: cell wall

## ACCESSIONS

Reference: A94961  
 Authors: Chen, J., Varner, J.E.  
 Journal: Proc. Natl. Acad. Sci. U.S.A. 110:257-262, 1994  
 Title: Isolation and characterization of cDNA clones for carrot  
 extension and a proline-rich 3xK protein

## CLASSIFICATION

Summary: #superfamily proline-rich protein  
 Search: 211  
 Query Match: 100.0%  
 Best Local Similarity: 100.0%  
 Matches: 6; Conservative: 6; Mismatches: 6; Indels: 6; Gaps: 6

DB: 184 VWRP 189  
 QY: 1 VWRP 6

## RESULT

Entry: J02260  
 Title: 3xK proline-rich cell wall protein - carrot  
 Organism: Formal name: Nicotiana glauca; common name: carrot  
 Date: 10-Mar-1994  
 Accession: J02260  
 Molecule type: DNA  
 Residues: 1-211  
 Cross-references: P01477; N01477; N01477-24  
 Experimental source: vat. Nates  
 Comment: This protein is a wall and is involved in cell  
 Classification: #superfamily proline-rich protein  
 Keywords: cell wall

## ACCESSIONS

Reference: J02260  
 Authors: Blenket, W.; Exler, T.; Suzuki, H.; Stover, J.; Henry,  
 M.L.  
 Journal: Plant Physiology (1993) 101:259-265  
 Title: Expression of carrot is linked to carrot storage root  
 formation and is induced by wounding and auxin treatment

SUMMARY #length 277 #molecular\_weight 47657 #checksum 5427

Query Match 100.0% Score 47. DB 2 Length 235  
 Best Local Similarity 100.0% Pred No 5.48e+00  
 Matches 5 Mismatches 0 Indels 0 Gaps 0

Db 208 VWRPPP 213  
 |||||  
 QY 1 VWRPPP 6

RESULT 4  
 ENTRY S29903 #type fragment  
 TITLE Cell wall peptidoglycan protein - English inv. (fragment)  
 ORGANISM #formal\_name Hodera helix #common\_name English inv  
 DATE 06-Jan-1995 #sequence\_revision 06 Jan 1995 #text\_change 09-Sep-1997

ACCESSIONS S29903  
 REFERENCE S29903  
 #authors Woot, H  
 #submission submitted to the EMBL Data Library September 1992  
 #accession S29903  
 #status preliminary  
 #molecule\_type RNA  
 #residues 1-99 #label WOO  
 #cross-references EMBL:Y28124, NID:919843, FID:342022  
 #length 99 #checksum 4474

Query Match 95.7% Score 45. DB 2 Length 99  
 Best Local Similarity 83.3% Pred No 5.48e+00  
 Matches 5 Mismatches 1 Indels 0 Gaps 0

Db 76 VWRPPP 81  
 |||||  
 QY 1 VWRPPP 6

RESULT 5  
 ENTRY I49681 #type complete  
 TITLE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.1.1.41)  
 ORGANISM mouse  
 DATE 10-Aug-1995 #sequence\_revision 10-Aug-1995 #text\_change 13-Mar-1997

ACCESSIONS I49681, I49646  
 REFERENCE I49681  
 #authors Welch, J.E., Schaffer, F.C., O'Brien, D.A., Eddy, P.M.  
 #journal PLoS. Med. (1997) 4:869-878  
 #title Expression of a glyceraldehyde 3-phosphate dehydrogenase gene specific to mouse spermatogenic cells.

#cross-references M0109237322  
 #accession I49681  
 #status preliminary  
 #molecule\_type mRNA  
 #residues 1-440 #label RFS  
 #cross-references GDB:403429, NID:919417, FID:497704

REFERENCE I49046  
 #authors Welch, J.E., Brown, P.E., O'Brien, D.A., Eddy, P.M.  
 #journal Dev. Genet. (1995) 16:179-189  
 #title Gene expression of a mouse glyceraldehyde 3-phosphate dehydrogenase gene (Gpdh-s) expressed in post-meiotic spermatogenic cells.

#cross-references M010923746  
 #accession I49046  
 #status preliminary, translated from SE/EMBL/DBJ  
 #molecule\_type DNA  
 #residues 1-32,95-42,117,14-440 #label R2  
 #cross-references EMBL:069644, NID:9457412, FID:9457413

GENETICS  
 #gene Gpdh-s  
 #introns 371, 114, 141, 181, 211, 251, 271, 311, 341, 371, 401, 431

CLASSIFICATION #superfamily glyceraldehyde 3-phosphate dehydrogenase

SUMMARY #length 440 #molecular\_weight 47657 #checksum 7527

Query Match 95.7% Score 45. DB 2 Length 440  
 Best Local Similarity 83.3% Pred No 5.48e+00  
 Matches 5 Mismatches 0 Indels 0 Gaps 0

Db 35 VWRPPP 40  
 |||||  
 QY 1 VWRPPP 6

RESULT 6  
 ENTRY I4112 #type complete  
 TITLE Probable indole-acetate beta-glucosyltransferase - Arabidopsis thaliana  
 ORGANISM #formal\_name Arabidopsis thaliana #common\_name mouse-ear cress  
 #accession D71419  
 #status preliminary  
 #molecule\_type RNA  
 #residues 1-99 #label WOO  
 #cross-references EMBL:Y28124, NID:919843, FID:342022  
 #length 99 #checksum 4474

ACCESSIONS D71419  
 REFERENCE A71400  
 #authors Dean, C., Bergkamp, P., Dirkse, W., Van Staveren, M., Zickema, W., Fries, L., Filday, F., Gibson, S.A., Patel, K., Murphy, G., Piffanelli, P., Wedler, H., Wedler, E., Wambutt, R., Weitzenecker, T., Pohl, T.M., Terryn, N., Giesen, J., Villarroel, P., De Clerck, E., Van Montagu, M., Lecharny, A., Aubry, S., Vandenbroucke, M., Tao, N., Ravanah, T., Hompl, S., Vetter, P., Fritsch, E.D., Piener, M., Schaeffer, M., Funk, B., Mueller-Auer, S., Silvey, M., James, K., Montfort, A., Pous, A., Rodriguez, P., Boucka, A., Vouklatov, P., Milioni, D., Hatzopoulos, P., Piravandi, E., Obermaier, B., Hilbert, H., Duesterhoft, A., Moeres, T., Jones, J.D.G., Eneva, T., Palme, K., Benes, V., Pechman, S., Ansoy, W., Cooke, R., Berger, C., Belseny, M., Voer, M., Volckaert, G., Mewes, H.W., Kistnerman, S., Schueller, C., Chelwatzis, N.

#journal Nature (1998) 391:485-488  
 #title Analysis of 1.8 Mb of contiguous sequence from chromosome 4 of Arabidopsis thaliana  
 #cross-references M010923113  
 #accession D71419  
 #status preliminary, nucleic acid sequence not shown; translation not shown

#molecule\_type DNA  
 #residues 1-484 #label BEV  
 #cross-references GDB:403429, NID:919417, FID:497704

GENETICS  
 #pos-position 4039-433845  
 #length 484 #molecular\_weight 53876 #checksum 3179

SUMMARY #length 440 #molecular\_weight 53876 #checksum 3179

Query Match 95.7% Score 45. DB 2 Length 440  
 Best Local Similarity 83.3% Pred No 5.48e+00  
 Matches 5 Mismatches 0 Indels 0 Gaps 0

Db 316 VWRPPP 321  
 |||||  
 QY 1 VWRPPP 6

Search completed: Fri Feb 25 13:21:59 2000  
 Job time : 52 secs.



Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

DB 35 VIRPPP 40  
 QY 1 VVRPPP 6  
 Search time: 0.03 Fri Feb 25 14:07:10 2000  
 Job time: 45 secs

```

QY 1 VVRPPP 6
RESULT 3
ID G3PT_MOUSE STANDARD: PRT: 440 AA.
AC Q6445: 1000000:
DE C: NOV-1997 (REL. 35, CREATED)
DI 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DI 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE, TESTIS-SPECIFIC
DE EC 1.2.1.12 (GAPDH).
GN GAPDS OR GAPD-S.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA, METAZOA, CHORDATA, VERTEBRATA, MAMMALIA, EUTHERIA.
OC RODENTIA: SCIUROGNATHI: MURIDAE: MURINAE: MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=TESTIS;
RX MEDLINE: 92222722
RA WELCH J.E., SCHATTE E.C., O'BRIEN D.A., EDDY E.M.:
RT "Expression of a glyceraldehyde 3-phosphate dehydrogenase gene
specific to mouse spermatogenic cells.";
RL RIGL. REPROD. 45:869-878(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR SWISS; TISSUE=TESTIS;
RX MEDLINE: 95254745.
RA WELCH J.E., BROWN P.R., O'BRIEN D.A., EDDY E.M.:
RT "Genomic organization of a mouse glyceraldehyde 3-phosphate
dehydrogenase gene (Gapd-s) expressed in post-meiotic spermatogenic
cells.";
RL DEV. GENET. 16:179-189(1995).
CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN REGULATING THE SWITCH
BETWEEN DIFFERENT PATHWAYS FOR ENERGY PRODUCTION DURING
SPERMIOGENESIS AND IN THE SPERMATOZOON.
CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + OXYPHOSPHATE
- NAD(+) -> 1,3-DIPHOSPHATEGLYCERATE + NADH.
CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: TESTIS-SPECIFIC.
CC -1- DEVELOPMENTAL STAGE: FIRST EXPRESSED AT DAY 20 IN POST-MEIOIC
GERM CELLS. LEVELS INCREASE UNTIL DAY 24 AND THEN REMAIN CONSTANT
DURING MATURITY.
CC This Swiss PDB entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed, changed, or for commercial
entities requires a license agreement. (See: http://www.isb-sib.ch)
or send an email to license@sib-sib.ch).
CC -----
DR ENRL: M60978: G567204: -
DR EMBL: U09964: G497413: -
DR MGD: MGI:95653: GAPDS.
DR PROSITE: PS00071: GAPDH; 1.
DR PFAM: PF00044: gapdh; 1.
DR KX GLYCOLYSIS; OXIDOPHOSPHATASE; NAD.
FT DOMAIN 21 41 CYS/PRO-RICH.
FT DOMAIN 54 73 POLY-PRO.
FT DOMAIN 84 100 POLY-PRO.
FT BINDING 256 256 GLYCERALDEHYDE 3-PHOSPHATE (BY
SIMILARITY).
FT ACT_SITE 263 283 ACTIVATES THIOL GROUP DURING CATALYSIS
(BY SIMILARITY).
FT CONFLICT 33 34 MISSING (IN REF. 2).
FT CONFLICT 43 43 L -> V (IN REF. 2).
SQ SEQUENCE 440 AA: 47657 MW: 40693.25 pI:5.02
Query Match 95.7% Score 45; DB 1; Length 440;
Best Local Similarity 83.3%; Pred. No. 1.42e-00;

```



Matches 5, Conservative 1, Mismatches 0, Indels 0, Gaps 0;

DB 52 VRRPP 67  
1 VRRPP 6

RESULT 4  
ID 018512 PRELIMINARY: PRI: 445 AA.  
AC 018512:  
DI 01-JAN-1998 (PRELIMEL 05, CREATED)  
DT 01-JAN-1998 (PRELIMEL 05, LAST SEQUENCE UPDATE)  
DI 01-NOV-1998 (PRELIMEL 08, LAST ANNOTATION UPDATE)  
DE 01-NOV-1998 (PRELIMEL 08, LAST ANNOTATION UPDATE)  
DE SEROTININ RECEPTOR PRECURSOR.  
OS CAENOHABDIIIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SPOTENENTIA; PHARITIA; PHARITIA;  
OC PHARITIA; PHARITIA; PHARITIA; PHARITIA; PHARITIA; PHARITIA;  
RN [1]  
RF SEQUENCE FROM N.A.  
RA OLDE B., MCCOMBIE R.W.;  
RL J. MOL. NEUROSOL. 6:0-0(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA DUBNICK M.J.  
F1 SPENTIER (1994) J. ENH. JENKINS, JER. CAIA BANKS  
C1 SUBCELLULAR LOCATION: INTERAL MEMBRANE FRILLO (BY SIMILARITY);  
OR ENBL. 015167, 02317845;  
DR PROSITE: PS00379; G-PROTEIN-RECEPTOR; 1  
DR PFAM: PF00001; 740-1; 1.  
FW JEFFERY JEFFERY, FEETIFF, TRANSMEMBRANE, MYOEPITIN  
SQ SEQUENCE 445 AA, 50141 MW, 457829 CRO32;

Query Match 95.7% Score 45; DB 5; Length 445;  
Best Local Similarity 83.3% Pred. No. 2,170-00;

Matches 5, Conservative 1, Mismatches 0, Indels 0, Gaps 0;

DB 253 VRRPP 218  
1 VRRPP 6

RESULT 5  
ID 023400 PRELIMINARY: PRI: 484 AA.  
AC 023400:  
DI 01-JAN-1998 (PRELIMEL 05, CREATED)  
DT 01-JAN-1998 (PRELIMEL 05, LAST SEQUENCE UPDATE)  
DI 01-NOV-1998 (PRELIMEL 08, LAST ANNOTATION UPDATE)  
DE INDOL-3-ACETATE BETA-GLUCOSYLTRANSFERASE.  
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).  
OC EUKARYOTA; VIRIDIPANTAE; STEPHENTHIA; EUPHYLLIPHYTA; FRACHIOPHYTA;  
OC EUPHYLLIPHYTES; SPERMATOPHYTES; MAGNOLIOPHYTES; EUDICOTYLEDONS; ROSIDAE;  
OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA REVAN M., STITKMA W., MURPHY G., WAMRITT R., POHL T., TERRY N.,  
RA KEIS M., KAVANAGH T., ENIAN K.D., RIEGER M., JAMES P.,  
RA FUGLOMENER P., HATZPOULOS P., OBERMAIER B., DUESERHOFT A.,  
RA JONES J., PALME K., ANSOFER W., PRISNY M., RANCOFT I., MEWIS H.W.,  
RA SCHUELLER C., CHALWATZIS N.,  
F1 SUBMITTED 1997-1997 TO ENBL, JENKINS, JER. CAIA BANKS.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU ARABIDOPSIS SEQUENCING PROJECT, ESSA;  
F1 SPENTIER (1994) J. ENH. JENKINS, JER. CAIA BANKS.  
OR ENBL. 027339, 0307472;  
DR PROSITE: PS00375; UDPGT; 1.  
DR PFAM: PF00261; UDPGT; 1.  
DR TRANSFERASE.  
SQ SEQUENCE 484 AA, 53876 MW, 8803788 CRO32;

Query Match 95.7% Score 45; DB 10; Length 484;  
Best Local Similarity 83.3% Pred. No. 2,170-00;

Matches 5, Conservative 1, Mismatches 0, Indels 0, Gaps 0;

DB 315 VRRPP 321  
1 VRRPP 6

RESULT 6  
ID 013024 PRELIMINARY: PRI: 873 AA.  
AC 013024:  
DI 01-JUL-1997 (PRELIMEL 04, CFFAIED)  
DT 01-JUL-1997 (PRELIMEL 04, LAST SEQUENCE UPDATE)  
DI 01-NOV-1998 (PRELIMEL 08, LAST ANNOTATION UPDATE)  
DE XL-INCENP.  
OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;  
OC MESOBATRACHIA; PIPIDAE; PIPIDAE; XENOPUS.  
RN [1]  
RF SEQUENCE FROM N.A.  
RX MEDLINE: 9727270.  
RA STRENNER P.T., DUSTIN K.D., MCCARRY T.J., KING R.W., KIANG J.,  
RA KIRSCHNER M.W.;  
F1 "Systematic identification of mitotic phosphoproteins";  
PL CURR. BIOL. 7:338-348(1997).  
OR ENBL. 015167, 02317845;  
F1 SPENTIER (1994) J. ENH. JENKINS, JER. CAIA BANKS

Query Match 95.7% Score 45; DB 13; Length 873;  
Best Local Similarity 83.3% Pred. No. 2,170-00;

Matches 5, Conservative 1, Mismatches 0, Indels 0, Gaps 0;

DB 445 VRRPP 451  
1 VRRPP 6

Search Completed: Fri Feb 25 13:29:14 2000  
Job time : 76 secs.